Carl Lamote

Less Successful Pathways Through Secondary School

KATHOLIEKE UNIVERSITEIT LEUVEN
FACULTEIT PSYCHOLOGIE EN PEDAGOGISCHE WETENSCHAPPEN
Onderzoekseenheid Onderwijskunde

Centrum voor Onderwijseffectiviteit en -evaluatie
Centrum voor Methodologie van het Pedagogisch Onderzoek

Less Successful Pathways Through Secondary School

Studies on grade retention and early school leaving

Proefschrift aangeboden tot
het verkrijgen van de graad van
Doctor in de Pedagogische Wetenschappen

door **Carl Lamote**

o.l.v. Prof. Dr. Jan Van Damme
Prof. Dr. Wim Van Den Noortgate

2013

KU LEUVEN

FACULTEIT PSYCHOLOGIE EN
PEDAGOGISCHE WETENSCHAPPEN

ONDERZOEKSEENHEID ONDERWIJSKUNDE
CENTRUM VOOR ONDERWIJSEFFECTIVITEIT EN –EVALUATIE
CENTRUM VOOR METHODOLOGIE VAN HET PEDAGOGISCH ONDERZOEK

Less successful pathways through secondary school studies on grade retention and early school leaving

Carl Lamote

Proefschrift aangeboden tot het verkrijgen van de
graad van Doctor in de Pedagogische Wetenschappen

Promotor: Prof. Dr. Jan Van Damme
Co-promotor: Prof. Dr. Wim Van Den Noortgate

2013

Carl Lamote, Less successful pathways through secondary school: studies on grade retention and early school leaving

Doctoral dissertation submitted to obtain the degree of Doctor in Educational Sciences, 2013.

Supervisor: Prof. Dr. Jan Van Damme
Co-supervisor: Prof. Dr. Wim Van Den Noortgate

In this dissertation, we focus on the predictors of early school leaving (ESL) and the effects of grade retention. Although we mention them separately, both educational phenomena are closely related, which will become clear in this dissertation. Over the years, both grade retention and ESL were the subject of many studies. These studies had some shortcomings, related to their methodological approach, as well as more theoretical shortcomings. In the present dissertation, we attempt to meet these shortcomings, by making use of a rich, longitudinal database (LOSO: Dutch acronym for Longitudinal Research in Secondary Education), in combination with the appropriate methodology. In the *first manuscript*, we explored the ESL-phenomenon in Flanders, and related this to student mobility. To investigate this, we applied a 'multilevel discrete-time survival' model, where we account for school changing. We compare different approaches of handling such school changers. From our results, it is clear that the standard, pragmatic solutions (can) have implications on the results and that working with cross-classified (or 'last school') models is the most appropriate way of handling student mobility. Next to that, this manuscript provided evidence for the relation between ESL and some (internationally prevalent) predictors of ESL in Flanders. From our results, it became clear that boys, students with a low initial achievement, students with a low socioeconomic status (SES), grade retainees, students in a school with a low SES and students in schools with a less positive student-teacher relation, are more likely to drop out. In the *second manuscript*, we focus on the development of emotional and behavioural engagement and the relation with ESL. By making use of a 'discrete-time survival mixture' model, we can distinguish several trajectories in the development of both engagement dimensions. The membership of the student of one of these trajectories is used to predict ESL. For behavioural engagement we found three groups of students: students starting at a high level of engagement and following a stable pattern, students starting at a high level but following a decreasing trend, and students starting at a low level of behavioural engagement. These last two groups are more likely to drop out, compared to the first group of students. A similar conclusion can be drawn for emotional engagement, where we found two groups: students starting at a high level and following a stable trajectory, and students starting at a low level of engagement. The latter group was, again, more likely to leave school prematurely. Students in the more 'problematic engagement' groups (i.e. groups with a higher likelihood of dropout) are mainly characterized by a lower SES, a history of grade retention and a lower achievement. Boys are also more likely to be member of these groups, and therefore have a higher likelihood of dropping out. In the *third manuscript*, we turn to one of the strongest predictor of ESL: grade retention. We mainly focus on the effect of grade retention in Grade 8, on achievement and academic self-concept (the choice for Grade 8 is mainly because of data-technical reasons). Before we could analyse the effect of grade retention, we had to match students on their propensity score (in this case: the propensity to be retained in Grade 8). After this matching, it became clear that grade retainees perform (the second time in Grade 8) at the same level as students with the same propensity to be retained in Grade 8, a year earlier. It that sense, it may seem that grade retainees do 'catch up', and the practice of grade retention is a good solution. However, the years after grade retention, the achievement of grade retainees decreases, compared to students with the same propensity to be retained but who got promoted instead. Concerning the effect of grade retention on academic self-concept, we can conclude that grade retention gives a boost to the academic self-concept in the year of retention, and that the level of academic self-concept remains at a (small but significant) higher level. In *manuscript 4*, we also analysed the effect of grade retention in secondary education, but our outcomes of interest were at the post-secondary level where we focus on enrolment and success in higher education. After a carefull matching on the propensity score, we find that students retained in any grade of secondary education are less likely to enrol in higher education. Concerning the effect of grade retention in secondary education on success in higher education, we find that being retained in the last three grades of secondary school has a significant negative effect on success in higher education. We conclude this dissertation by pointing at some limitations of our studies, and discuss some practical limitations.

Carl Lamote, Minder succesvolle wegen doorheen het secundair onderwijs: studies over zittenblijven en vroegtijdig schoolverlaten

Proefschrift aangeboden tot het verkrijgen van de graad van Doctor in de Pedagogische Wetenschappen, 2013.

Promotor: Prof. Dr. Jan Van Damme
Co-promotor: Prof. Dr. Wim Van Den Noortgate

In dit proefschrift focussen we op de voorspellers van vroegtijdig schoolverlaten (vsv) en op de effecten van zittenblijven. Hoewel we deze apart vermelden, zijn beide onderwijskundige fenomenen sterk met elkaar verbonden, zoals zal blijken uit dit proefschrift. Over de jaren heen was zowel vsv als zittenblijven het onderwerp van een groot aantal studies. Deze studies vertoonden echter steeds beperkingen van verschillende aard: methodologisch maar evenzeer inhoudelijk. In het voorliggende proefschrift trachten we hieraan tegemoet te komen, door gebruik te maken van een rijke longitudinale databank (LOSO; Longitudinaal Onderzoek Secundair Onderwijs) in combinatie met de aangewezen methodologie. Het *eerste manuscript* verkent het probleem van vsv in Vlaanderen, met speciale aandacht voor schoolveranderaars. Daartoe gebruiken we een 'multilevel discrete-time survival' model, waarbij we rekening houden met schoolveranderaars. We vergelijken daarbij verschillende aanpakken om om te gaan met zulke schoolveranderaars. Uit onze resultaten blijkt dat de meer courante en pragmatische aanpakken implicaties hebben voor de resultaten en dat het werken met kruisclassificaties (of met de laatste school als classificatie) het meest aangewezen is. Bijkomend biedt dit manuscript ook evidentie voor de relatie tussen vsv en enkele (internationaal courante) voorspellers in Vlaanderen. Zo blijken jongens, leerlingen met een lage aanvangsprestatie, leerlingen met een lage sociaaleconomische achtergrond (SES), zittenblijvers, leerlingen in een school met een lagere sociaaleconomische compositie en leerlingen in een school met een gemiddeld minder goede leerling-leerkracht relaties een hogere kans te hebben op vsv. In het *tweede manuscript* focussen we op het verloop van de schoolse betrokkenheid van leerlingen op de kans op vsv, waarbij we aandacht hebben voor het verloop van zowel gedragsmatige als emotionele betrokkenheid. Door gebruik te maken van een 'discrete-time survival mixture' model, kunnen we voor beide dimensies van betrokkenheid verschillende groepen qua verloop van betrokkenheid onderscheiden. Het lidmaatschap van deze groepen gebruiken we vervolgens om de kans op vsv te bepalen. Zo blijken er voor gedragsmatige betrokkenheid drie groepen leerlingen te zijn: leerlingen met een hoog en stabiel verloop van betrokkenheid, leerlingen die hoog starten maar snel een dalend verloop kennen en leerlingen die het secundair onderwijs reeds aanvatten met een laag niveau van betrokkenheid. Deze laatste twee groepen kennen een sterk verhoogde kans op vsv. Voor emotionele betrokkenheid geldt een gelijkaardige conclusie: daar onderscheiden we twee groepen, die vooral verschillen in aanvangsniveau. Zo vinden we een groep leerlingen die hoog start en een stabiel verloop kent, en een groep leerlingen die laag start en een dalend verloop kent. Ook deze laatste groep kent een sterk verhoogde kans op vsv. Leerlingen in de 'problematische-betrokkenheids' groepen (d.i. de groepen die eveneens een verhoogde kans op vsv kennen) worden vooral gekenmerkt door een lagere SES, een verleden van zittenblijven en een lagere prestatie. Bovendien hebben jongens meer kans om in deze groepen te zitten, en daardoor een verhoogde kans op vsv. In het *derde manuscript* focussen we op één sterke voorspeller van vsv, nl. zittenblijven. Daarbij hebben we vooral aandacht voor het effect van zittenblijven in het tweede jaar van het secundair onderwijs (de keuze voor dit jaar is voornamelijk gebaseerd op data-technische redenen) op de verdere prestaties en het academisch zelfconcept. Vooraleer we echter het effect van zittenblijven op de uitkomstvariabelen konden analyseren, dienden we de leerlingen te matchen op basis van een propensity score (in dit geval: de kans op zittenblijven). Na deze matching bleek dat zittenblijvers in het jaar van zittenblijven gelijk presteren als leerlingen een jaar eerder in datzelfde leerjaar. In die zin lijkt zittenblijven een zinvolle praktijk. De jaren na het zittenblijven zien we echter de prestaties van deze zittenblijvers dalen, in vergelijking met leerlingen met een gelijkaardig profiel die wél naar een volgend leerjaar mochten. Voor wat betreft het academisch zelfconcept van de leerlingen, zien we in het jaar van zittenblijven een verhoogd academisch zelfconcept in vergelijking met leerlingen die een jaar eerder eenzelfde kans op zittenblijven hadden, maar wél naar het volgend leerjaar gingen. Dit verhoogde academisch zelfconcept houdt redelijk stand naarmate deze leerlingen vorderen doorheen het secundair onderwijs. In manuscript 4 kijken we eveneens naar de effecten van zittenblijven, maar is de uitkomstvariabele op post-secundair niveau en focussen we deelname en succes in hoger onderwijs. Na zorgvuldige matching op basis van de propensity score, blijkt dat leerlingen die in eender welk leerjaar van het secundair onderwijs bleven zitten, een significant lagere kans hadden op deelname aan hoger onderwijs. Wat betreft het succes in hoger onderwijs, blijkt er enkele een significant negatief effect te zijn van zittenblijven in de laatste drie jaar van het secundair onderwijs. We eindigen dit doctoraat met een overzicht van de limitaties en implicaties voor de praktijk.

DANKWOORD

Bij de start van mijn doctoraat, heb ik lang met het gevoel rondgelopen dat ik in de meest eenzame job ter wereld was terecht gekomen. Je kon vier dagen dood achter je bureau zitten vooraleer iemand iets zou opmerken. Althans zo leek het toch. Toch bleek gaandeweg dat dit doctoraat allerminst een eenzame job was/is, maar het sluitstuk van vier jaar interessante discussies, besprekingen, nalezen van teksten en ijsjes/taartjes eten. Ik ben daarvoor heel wat mensen uitermate dankbaar en probeer ze hier kort even voor te bedanken. De enkele regels die ze hier krijgen zullen echter nooit in verhouding staan tot wat ze voor mij en dit doctoraat hebben betekend. Maar toch een poging…

In de eerste plaats wil ik Jan en Wim bedanken. Jan, bedankt om tijdens het sollicitatiegesprek meer in mij te zien dan een jongeman naarstig op zoek naar een job en bedankt om mij na afloop van het gesprek het legendarische antwoord ‘we wagen het er op’ te sturen. Bedankt om me de kans te geven om aan het CO&E te werken, maar vooral bedankt voor de vrijheid in keuzes die ik kreeg. Wim, bedankt voor alle feedback en bijsturing. Je deed alsof domme vragen écht niet bestonden, hoewel ik er ongetwijfeld heel wat heb gesteld. Steeds stond er een antwoord klaar en ik wil je daarvoor bedanken.

Het CO&E leek soms een beetje op een duiventil, maar toch waren er heel wat collega’s waar ik steeds op kon rekenen: Sara, Maarten, Tinneke, Jerissa en Ine. Sara, ik denk dat we heel wat lief en leed hebben gedeeld als bureaugenoten en dat we elkaars ‘kleine kantjes’ hebben leren kennen. Bedankt dat ik steeds mijn jas aan de rechterkapstok mocht hangen, dat betekende heel veel voor mij. Bedankt ook voor de feedback op mijn eerste twee papers, het zijn tot hiertoe de enige die gepubliceerd zijn geraakt. Maarten, steeds stond je klaar met interessante inzichten, die mijn derde en vierde paper heel wat stappen vooruit hebben geholpen. Ik zal je echter steeds blijven herinneren om de (verhalen over teveel) pintjes, de Westvleteren-uitjes, het doorzakken in Cádiz en het extreem onderuitgezakt zitten in een bureaustoel. Ik wens je het allerbeste op je lange reis en een behouden thuiskomst. Tinneke, de uitspraak die me van jou de laatste maanden steeds zal bijblijven is: ‘mag ik eens komen zagen’. Wees gerust, in 73,2% van de gevallen

was het geen gezaag, maar waren het doctoraat-gerelateerde problemen waarmee ik evengoed worstelde. Het samen afsluiten van het doctoraat gaf een beetje het gevoel van ‘partner in crime’ en maakte dat ik steeds mijn vordering kon aftoetsen tegen die van jou (en besefte dat ik moest doorwerken). Jerissa, als collega-Antwerpenaar (die tijdens presentaties de Antwerpse ‘a’ en ‘i’ vaak evenmin kon onderdrukken), ben ik vooral jaloers op je (voor de buitenwereld) zorgeloze manier van leven. Hoewel kwatongen het ‘verwaaid’ durven noemen, ken ik niemand die met een meer open blik situaties bekijkt en zich steeds weer laat verrassen door iets nieuws. Een eigenschap die ik hopelijk ook bij mezelf ontdek en die ik dan zal koesteren. Ine, door jou aan het werk te zien, besef ik wat voor een titanenwerk een data-verzameling is. Hoewel zo’n werk blijkbaar niet steeds geheel foutloos verloopt, denk ik toch dat je apetrots mag zijn op wat de SiBO-databank nu is. Lijken zullen er altijd wel uit de kast vallen, maar ik beloof dat we het je dan niet gaan vertellen. En ik beloof ook dat ik nooit meer ga lachen met orthopedagogen.

Uiteraard bestond CO&E nog uit heel wat andere mensen, die van dichtbij of van verder weg betrokken waren bij dit doctoraat. Daarom eveneens een welgemeende dank aan Machteld, Gudrun, Kim, Sofie, Barbara, Chloé, Bieke, Hongqiang, Bo, Henri, Catharina en Eef. Bedankt voor eender welke input die jullie hebben geleverd, hetzij inhoudelijk, hetzij de ontspanning tijdens één of andere veel te lang uitgelopen lunchpauze.

Charlotte, als mijn bureaugenote het laatste halfjaar, krijg jij een aparte paragraaf. Ik denk dat je een zwempartij in heel woelig water achter de rug hebt, maar ik hoop dat je op CO&E in een rustige en gemoedelijke sfeer je doctoraat kan afwerken. Ik wens je daarbij veel geluk, moed en doorzetting.

David, bedankt om mij te vernoemen in jouw dankwoord; dat verdient dan ook een plaatsje in mijn dankwoord.

Katrien, misschien ietwat verrassend, maar ook voor jou een woordje van dank. Mij aanwerven nog vooraleer ik m’n doctoraat had, zorgde ervoor dat ik dit doctoraat met een gerust gemoed kon afwerken. Onbewust ontnam je mij zo een hele last van de schouders. Oh ja, zou ik de eerste week op mijn nieuwe job al een weekje verlof mogen nemen?

De laatste pagina is voorbehouden voor de mensen die al jaren dicht bij me staan. Paul, Gonda, Jonas, Saskia, Kirsten en Xavier, bedankt om me mee op te nemen in de clan en om de vele momenten samen die ervoor zorgden dat ik dit doctoraat even kon vergeten.

Steven en Leen, gewoon bedankt om mijn broer (en schoonzus) te zijn. Bedankt voor de afleiding de afgelopen vier jaar en bedankt om een goede peter van Mathijs en meter van Janne te zijn.

Moeke en papa, bedankt om al die jaren in mij te geloven, ook wanneer ik dat zelf niet meer deed. Bedankt voor alle ondersteuning, zowel logistiek, financieel als moreel. Zonder jullie waren – zeker de eerste jaren in hoger onderwijs – gewoonweg niet mogelijk geweest. Ik hoop dat dit doctoraat jullie met trots vervult, het is alles wat ik in me had.

En nu…

Hoe bedank je iemand die dag in, dag uit voor je klaarstaat? Hoe bedank je iemand die steeds een luisterend en oprecht geïnteresseerd oor aanbiedt, ook al weet je dat je ongelooflijk aan het zagen bent? Sofie, bedankt voor alles. Bedankt om er steeds voor mij te zijn. Bedankt om klaar te staan op momenten dat het moeilijk ging, maar ook bedankt om samen met mij de ‘kleine overwinningen’ te vieren, alsof het ook de jouwe waren. Ik kan me geen betere partner inbeelden, en geen betere moeder voor onze twee mooiste verwezenlijkingen…

Mathijs, Janne,

Dankzij jullie besef ik dat dit doctoraat (en werk in het algemeen) slechts één aspect is van het leven. Jullie bestaan doet me beseffen dat een mens niet veel nodig heeft om zielsgelukkig te zijn.

Bedankt om me dat te doen beseffen…

Carl Lamote

Leuven, 30 augustus 2013

TABLE OF CONTENT

Chapter 1
Introduction

EARLY SCHOOL LEAVING: FRAMING THE PROBLEM

For many years, early school leaving was the topic of multiple studies (see e.g. Alexander, Entwisle, & Horsey, 1997; Finn, 1989; Jimerson, 2001; Rumberger, 1995), in which the focus was not only on trying to explain the dropout phenomenon from different viewpoints, but also on formulating a number of prevention or intervention strategies, based on the explanatory variables of dropout. The growing attention for this dropout-phenomenon can be explained by the major consequences of leaving secondary education without a diploma or qualification. In an age where diplomas are continuously gaining in importance, leaving school without a basic qualification implies a strong decrease in future employment opportunities. One look at unemployment statistics or other indicators related to success on the labour market makes it clear: graduating with a basic qualification (i.e. generally speaking: a diploma of secondary school, for more details see further) lowers the chances of (long-term) unemployment and guarantees better outcomes (on earnings, health, etc.) in later life. The impact of a qualification on unemployment is made explicit in Figure 1, in which the unemployment rate (in 2012) by International Standard Classification of Education (ISCED) level is depicted for the EU15 countries (EUROSTAT, 2013). This figure clearly shows that reaching an ISCED level 3 or 4, results in lower unemployment, compared to students that only reached at most ISCED level 2. Students only reaching this low level of education are usually referred to as 'early school leavers' or 'dropouts'[i] and will be the focus of this dissertation.

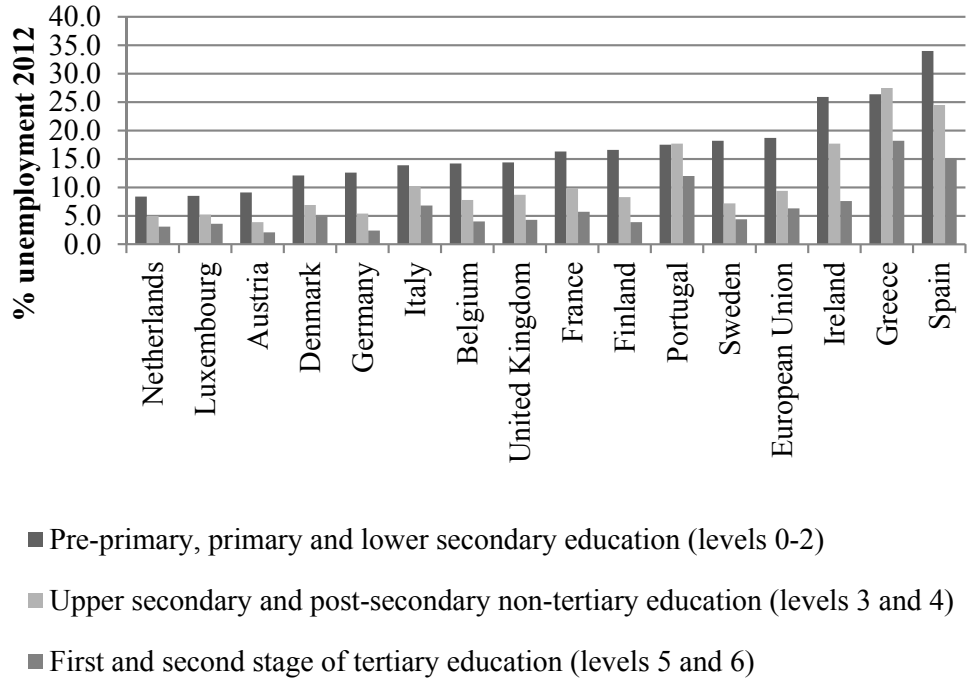


Figure 1: unemployment rate of EU15-citizens between the age of 15-64, by ISCED-level

A qualification: how does it make a difference?

A considerable amount of research and theories have been published on the importance of a diploma for both the employer and employee. These theories include, among others, human capital theories (which considers schooling as a way of obtaining necessary skills and abilities for the labour market), screening and signalling theories (where employers screen the future employees in search of *the right man for the right job* and where a diploma acts as a signal for these employers), and credentialist theories (schooling leads to socio-economic success because schooled employees have more access to elite positions) (Bills, 2003). Although neither of these theories can fully explain the subtle relations between schooling (and diploma) and job opportunities, most of the studies agree on the importance of a qualification for future job opportunities and confirm the negative effect of leaving school without a qualification. Besides the importance of a qualification for the individual, Hanushek (2009) stresses the importance of a qualification for the society as a whole: human capital affects the improvements in productivity and national income, and thereby accelerates economic growth. The

more educated people, the more knowledge produced in a society, which has a strong impact on the economic growth of a country. Psacharopoulos (2007) provides a more direct link between qualifications and economic growth by pointing out the relation between higher unemployment rates amongst unqualified people and the higher unemployment payments paid by the government.

Because of the importance of a qualification, studies on dropout remain of high importance in educational research, and a growing body of studies identified several predictors for dropout. In his research, Rumberger (2011) groups these predictors in two large sets: predictors at the individual level on the one hand, and predictors at the institutional level on the other hand.

Individual perspective

The individual perspective refers to a broad set of student characteristics associated with the decision to withdraw from school. In the following paragraphs, we will discuss three relevant (for this dissertation) and frequently cited predictors of dropout: academic achievement, gender and engagement.

Achievement

A strong predictor of dropout is the academic achievement of the student, as reflected in test scores. The argument that test scores predict dropout or graduation in secondary education may seem very rational and even trivial, since these test scores should reflect the competence of the student. What is of interest in relation to these test scores, is the strength of the predictive power. This predictor is not limited to academic achievement in secondary education; even the academic achievement in primary education is a strong predictor for dropout in secondary education. As Alexander, Entwisle and Kabbani (2001) conclude, even achievement in the first grade predict future dropout. Children whose test scores in first grade were in the 'A-B' (=highest) range have a much lower probability of dropping out, compared to children in the 'D-F' range (=lowest). Receiving lower grades in secondary education has an additional effect on the risk for dropout (although not every study is unanimous if it comes to the size of this effect on dropout). Bowers (2010) combined test scores from primary education and secondary education

starting from grade 1 up until grade 12, with mean grade point averages (GPA) for every grade. Students in the lowest GPA category had an overall higher risk of dropping out in every grade, starting from grade 6, and especially in grade 8 (30% of these low performing students dropped out) and grade 11 (45% dropped out!) (Bowers, 2010).

Gender

One of the most commonly considered student characteristic predicting dropout, is gender. Although regularly approached as an important predictor, the effect of gender on dropout was not always unequivocal: a first set of studies found that boys are more likely to leave school (e.g. Alexander et al., 2001; Marks, 2007), whereas a second set concluded an opposite effect with girls having higher chances to leave school (Croninger & Lee, 2001; Goldschmidt & Wang, 1999) and a third set of studies found no effect of gender on dropout (Lee & Burkham, 2003; Hickman, Bartholomew, Mathwig, & Heinrich, 2008). Apart of these main effects of gender, some studies concluded that gender affects dropout in interaction with the socio-economic status or ethnicity of the student (Luyten, Bosker, Dekkers, & Derks, 2003). This is comparable to what Rumberger (2011) concludes in his review: the reason of the mixed findings of the effect of gender is because of the relation with other variables. When a researcher only controls for background characteristics, it usually results in lower dropout rates for females (or no significant differences), whilst studies in which a researcher controls for attitudes, behaviours and performance, show higher dropout rates for females.

Engagement

Dropping out of secondary education is not only a matter of gender or low school performance, but can be an expression of a more subtle and underlying process of reduced school engagement. School engagement has been included in many dropout theories, but is not very well integrated in empirical research (see e.g. Archambault, Janosz, Fallu, & Pagani, 2009; Finn & Zimmer, 2012; Ream & Rumberger, 2008). When engagement is included in empirical research, it is often limited to a single measurement of engagement. However, a single measurement is not in accordance to the long-term process towards dropout. As Rumberger (1987) argues, dropping

out is the final point of a long-term process of disengagement, which often begins in primary school. According to this view, one should take into account this longitudinal character and implement this indicator as time-varying over different grades. Next, the definition of this concept seems to be ambiguous. In past research, this school engagement has been referred to in various ways, resulting in various measurements. In their review, Fredricks, Blumenfeld and Paris (2004) give an overview of the current knowledge on and definitions of school engagement, with special attention to different types of school engagement and the effects on student outcomes. They distinguish three types: (1) behavioural engagement, (2) emotional engagement and (3) cognitive engagement. Behavioural engagement focuses on the participation of the student in academic and social activities at school. Emotional engagement refers to the reactions to teachers, peers and the school in general. These reactions can be demonstrated by enthusiasm, optimism, curiosity and interest during an activity (Klem & Connel, 2004), and can be measured by work orientation, student-teachers relations and orientation towards the school. Cognitive engagement refers to the effort a student is willing to provide for schooling, which often depends on how students value schooling.

Institutional perspective

Most of the previously mentioned research was conducted only considering the student characteristics, which gives the impression that dropout is solely caused by the student. However, also the institutional settings in which students live can predict future dropout. Regarding this institutional perspective, we consider the family background of the student and the characteristics of the schools that students attend. The family background is usually operationalized by the socio-economic status (SES) of the student. Even when controlled for other individual characteristics, students with a low SES background have a higher dropout probability, compared to pupils with a high(er) SES (Alexander et al., 2001).

Concerning the second group of institutional variables, the characteristics of the school that students attend research is much more limited as only a few studies integrated school characteristics in a multilevel model (Goldschmidt & Wang, 1999; Lee & Burkam, 2003; Luyten, Bosker, Dekkers, & Derks, 2003;

Rumberger, 1995). Luyten et al. (2003) followed a cohort of 4,448 students in 133 schools, who started secondary education in 1989, for 5 years gathering information with regard to the position in the educational system each year. Several variables on school level concerning the school context, organization, opinions and classroom teaching were taken into account in a multilevel logistic regression. Only the percentage of students that moved on to higher education in the past, explained a substantial amount of the school variance in dropout rate. Lee and Burkam (2003) focused on school structure and social organization and concluded that students were less likely to drop out in schools where relationships between teachers and students are more positive. Rumberger and Thomas (2000) concluded that high-SES schools had 40% lower dropout rates than average SES schools, whereas low-SES schools had 60% higher dropout rates than an average SES school.

Several other institutional variables are related to early school leaving, but we only selected the variables of interest for the subsequent chapters. For a more complete overview of institutional variables affecting dropout, see e.g. De Witte, Nicaise, Lavrijsen, Van Landeghem, Lamote, and Van Damme (2013) or Rumberger (2011).

GRADE RETENTION

In the previous paragraphs, we deliberately left out one of the main predictors of early school leaving: grade retention (e.g. Jimerson, Anderson, & Whipple, 2002). We left it out the previous discussion, because within the scope of this dissertation, we believe it deserves a more thorough discussion. In this paragraph, we discuss the effects of grade retention on achievement and non-cognitive outcomes, but in the next paragraph, we zoom in on the relation between grade retention and early school leaving.

Yearly, a substantive amount of students repeats a grade in primary or secondary education. In the Flemish educational system, (based on the LOSO-data, see further) only 58,1% of the students complete secondary education in 6 years[ii]; 18,4% in 7 years and 6,7% in more than 7 years (and 16.8% leaves school without a

diploma, with or without delay) (Van Damme, Meyer, De Troy, & Mertens, 2001). The non-promotion of some students is not without controversy, since several studies points out the negative effects of this non-promotion. Traditionally, these studies were conducted at the primary school level and only few studies accounted for grade retention in secondary schools (e.g. Kloosterman & De Graaf, 2010; Rodney, Crafter, Rodney, & Mupier, 1999; Spruyt, Laurijssen, & Van Dorsselaer, 2009). The dependent variables of the studies on grade retention can be divided into two broad categories: academic achievement on the one hand and psychosocial functioning on the other hand.

Past research on the effects of grade retention on academic achievement yielded different results depending on whether the effect was considered on the short term or on the long term. On the short term, during the year of retention, mainly positive effects were found: during the year of retention, non-promoted students performed better (Alexander, Entwisle, & Dauber, 2003; Karweit, 1999; Wu, West, & Hughes, 2008) compared to their new classmates, at least at the beginning of the school year. During the repeating year, the achievement of the non-promoted students relatively declines, and by the end of the repeating school year they are behind of their classmates (Bonvin, Bless, & Schuepbach, 2008). This result is a replication of previous studies (see e.g. Alexander et al., 1994; Wu et al., 2008) and is replicated in subsequent studies, with minor differences depending on the course subject. In general, studies considering the long-term effects of grade retention conclude that the positive effect in the (beginning of) the repeated year usually diminishes (Bonvin et al. 2008; Karweit, 1999), disappears (Jimerson, Carlson, Rotert, Egeland, Sroufe, 1997) or even shifts to a negative effect (Alexander et al., 2003). This negative effect on the long term is one of the main arguments for the opponents of grade retention.

However, a lot of these effects were recently tempered by the meta-analysis of Allen, Chen, Willson and Hughes (2009). They concluded that studies which reported the most negative effects of grade retention on academic achievement, were often studies with serious methodological problems. Therefore, they conducted a meta-analysis in which they only relied on methodologically sound studies (published before 2007), and they concluded that those studies only

show a small negative or even no effect of grade retention on academic achievement. Does this mean that grade retention is beneficial for students. Allen et al. (2009) correctly noted that "a finding of 'no significant difference' for retention on achievement calls into question the educational benefits of grade retention policies" (p. 493).

Regarding the psychosocial outcomes of grade retention, two theories are commonly cited: the social comparison theory of Festinger (1954) and the labelling theory (Becker, 1963). The social comparison theory is most often used by teachers and parents in support for grade retention. When repeating a grade because of low achievement, a student will be exposed to the subject matter of that particular grade for a second time, but will find himself in a classroom with students who never received this subject matter, giving the retained student a head start. Consequently, the retained student will feel more competent in this matter, compared to his new classmates. As a result, this head start provides the retained student with a higher level of perceived competence and a higher level of academic interest (Hong & Yu, 2008). During this extra year, the student will regain confidence as he can keep up again. However, this positive effect on academic interest seems to be of short duration: by the end of the school year, this positive effect was diminished (Bonvin et al., 2008). In contrast to the social comparison theory, the labelling theory (Becker, 1963) assumes that students retained in grade are given the label of 'stupid' students or 'failures' by their new, younger, classmates. This label of 'repeater' can potentially lead to a decline in school engagement and self-esteem. The effect of grade retention on the social development of students seems to be related with the moment of grade retention: early grade retention seems to have a smaller and less negative impact on a student's self-concept, compared to later grade retention and as such, students retained in kindergarten do not suffer from the 'retainee' label as much as students retained in secondary education (Shepard, 1989).

GRADE RETENTION AND DROPOUT

In several studies, the link between grade retention and dropout was underlined. Roderick (1994) concluded that, even after controlling for background variables and school performance, students who repeated a grade were more likely to dropout. In their systematic review, Jimerson et al. (2002) confirmed this effect of grade retention because all studies considered in the review identified grade retention as (one of) the most powerful predictor(s) of dropout. Different theories tried to explain this effect of grade retention. According to the frustration self-esteem theory (Finn, 1989), grade retention will lead to a decline in the student's self-esteem. This decline in self-esteem will, in turn, lead to withdrawal from the context seen as responsible for this decline. In the same study, Finn also introduced the participation – identification model. Students who identify with their school feel that they are part of the school and as a result, these students are more likely to participate in the curriculum and in extracurricular activities. This identification with the school is mediated by successful school outcomes; students with lower school outcomes (e.g. grade retention) tend to show less identification with their school. This lower level of identification will influence the participation in school activities and will result in higher levels of truancy. This, in turn, will lead to even lower levels of academic achievement and these students enter a vicious circle resulting in dropout. A third theory on the relation between dropout and grade retention refers to social capital. As a result of grade retention, social bonds with peers are broken. These students are separated from their former classmates and need to install new relations with their new classmates. This is, however, not easy since retained students tend to carry a label of 'failure' (Hong & Yu, 2008).

These three theories were tested by Stearns, Moller, Blau, and Potochnick (2007). At first, they confirmed that retained students are more likely to dropout, compared to continuously promoted students, even after controlling for their educational background. In testing the frustration – self-esteem model, these authors implemented self-concept as a measure for self-esteem. This variable, however, did not explain the probability of dropping out and retained students were still more likely to dropout. Based on these results, Stearns et al. concluded that the frustration – self-esteem model does not explain the relation between grade

retention and dropout. The participation and identification aspect in the second model of Finn (1989) were operationalized by two factors: student-participation (representing students' participation in extracurricular activities) and student-preparedness (representing students' attendance and preparedness for class). Both variables had a significant effect on dropout, while the effect of grade retention also remained significant. The participation – identification model explains a part of the decision to dropout, but cannot account for the relation between grade retention and dropout. The third and last model focused on the social capital of the student and analysed the parent-school and parent-student connectivity, bonds with teachers, the popularity with peers and school changes. This social-capital based model yielded similar conclusions as the participation – identification model: although several factors of the model explained the probability of dropping out (especially parent-student connectivity, lack of bonds with teachers and school changes), this social-capital based model could not explain the relation between grade retention and dropout. Based on these results, Stearns et al. conclude that existing (and commonly cited) models cannot fully explain the link between grade retention and dropout.

RESEARCH OBJECTIVES

In the previous sections, we provided an overview of the current knowledge on the matter of early school leaving and grade retention. Although the reader may get the impression that the knowledge on these phenomena is almost 'complete', there are some gaps to fill, starting with shortcomings of previous studies. These shortcomings will be the starting point of the next four chapters, and will be discussed thoroughly in these chapters. However, to give a general idea of these shortcomings and the subsequent chapters, we briefly discuss the problem statements and research objectives of the chapters in the following paragraphs.

Chapter 2

The starting point of this chapter is twofold: first, it will give an overview of the predictors of early school leaving in Flanders. Second, and more important, this chapter will treat some methodological aspects of analysing the phenomenon of early school leaving. Traditionally, dropping out is analysed as an event that takes

place at one point in time (usually at the end of compulsory education), on an individual level. Only few studies took into account the fact that students are nested within schools, and that school level factors may explain a part of the early school leaving phenomenon. In addition to ignoring the school level, most of the studies also ignored the fact that dropping out is only the endpoint of a long process (Rumberger, 2001), and that students can drop out at different points in time. Hence, analyses that do not account for this different moments of dropout are inappropriate because this important 'time' aspect is not fully covered. After all, Bowers (2010) concluded that time-varying variables are more important than time-invariant variables. In order to cover these two drawbacks (hierarchical nature and longitudinal character), one would suggest to use a longitudinal multilevel model. Although this seems appropriate, it treats the school as 'fixed', while in Flanders, changing schools is a very common practice, especially with (future) early school leavers. So, one can say that the school – and the corresponding school characteristics – are also 'time-varying' and should be handled in the appropriate way. To come towards all these critiques and shortcomings, we will model dropout in Flanders in a multilevel discrete-time survival framework, where we will compare a multiple membership approach with a cross classified approach (two approaches that explicitly take into account school changes) and other, more common, approaches.

Chapter 3

The central idea of this chapter builds on the statement of Rumberger (1987), who argues that early school leaving might better be viewed as a process of disengagement, rather than as an isolated action at a certain point in time. Two aspects of this statement were relevant for this chapter: the 'disengagement' component, and the 'process' component. Putting the concept of (dis)engagement at a central position in explaining dropout is more relevant than trying to explain dropout by only taking 'fixed factors' (e.g. gender) into account. Engagement as a central concept in explaining early school leaving is interesting, because this engagement is seen as malleable in a classroom setting, while the effect of e.g. gender or socioeconomic status is not. In their review of engagement, Fredricks et al. (2004) concluded that dropout research that included engagement in its analysis

usually only focussed on the behavioural aspect of engagement, while engagement is a multidimensional concept with, next to behavioural engagement, also an emotional and cognitive dimension. Fredricks et al. (2004) also noted that the gradualness of engagement in relation to early school leaving was often ignored, while "longitudinal research that explores the mediating processes between behavioural and emotional disengagement and dropping out is critical for intervention efforts" (p.72). This refers to the second component of Rumberger's statement: the 'process' component. Recently, some studies included engagement in a multidimensional way and combined this with a longitudinal approach (Archambault et al., 2009; Finn & Zimmer, 2012; Janosz, Archambault, Morizot, & Pagani, 2008). Still, these studies have to deal with some critique: or they only took two time points into account, which makes it difficult to study 'growth' (Finn & Zimmer, 2012), or they ignored the timing of dropout and the effect of gradual (dis)engagement on this timing (Archambault et al., 2009; Janosz et al., 2008). Therefore, we will conduct a discrete-time survival mixture analysis in which we will model the development of engagement and the survival of students simultaneously. This way, we will be able to formulate conclusions on the timing of dropout, predicted by the development of engagement. Moreover, since we include a 'mixture' aspect in our analysis, we will be able to capture subgroups with different growth patterns of engagement and to use these subgroups to find out if there are differences in timing of dropout. Concerning the multidimensionality of engagement, we will estimate the effect of (dis)engagement on early school leaving for behavioural and emotional engagement separately.

Chapter 4

Holding students back when they do not reach the required attainment level of a certain grade is common practice in several countries, including Flanders. Although this practice seems innocent and even positive for students (they receive 'the gift of time' to catch up with the material they did not (yet) master), there is a load of studies indicating the opposite, depending on the outcome under consideration (see previous section in this introduction). The majority of these studies only considered the effect of grade retention in primary education on achievement or non-cognitive outcomes in primary education, and only a minority looked at the effect of grade

retention in primary school on achievement in secondary education. In this chapter, we were interested in the effect of grade retention in secondary education, on different outcomes in secondary education. As far as we know, only two studies (with a sound methodology) focussed on this educational level (Ehmke, Drechsel, & Carstensen, 2010; Uysal, 2010). In our study, we will follow the methodological guidelines of Allen et al. (2009) and adopt a quasi-experimental approach where we construct two groups: students retained in Grade 8 and a group of comparable promoted students. By making use of a propensity score matching, we will be able to define these two groups, and to match retained students with promoted students who had a similar chance to retain (based on a large set of background variables) but were promoted instead. In a following step, we will include these matched students in two different growth curve analyses: a growth curve analysis where we focus on the effect grade retention on language achievement and a growth curve analysis where we focus on the effect of grade retention on academic self-concept.

Chapter 5

In this chapter, we will also focus on the effect of grade retention, but compared to the previous chapter, we are now more interested in the long-term effects of grade retention in secondary school. As long-term effect, we will consider different outcomes related to post-secondary education. The starting point for this chapter, are three previously published studies (Fine & Davis, 2003; Ou & Reynolds, 2010; Pustjens, Van de gaer, Van Damme, & Onghena, 2004) and the shortcomings of these studies. The shortcomings of these studies were threefold: (1) they only examined effects of grade retention in primary education or in the first grades of secondary education, and thereby ignored the effect of grade retention at the end of secondary education; (Fine & Davis, 2003; Ou & Reynolds, 2010) (2) grade retainees were not matched with comparable but normally promoted students, so it was not possible to differentiate between the effects of grade retention and student background characteristics (Fine & Davis, 2003; Pustjens, Van de gaer, Van Damme, & Onghena, 2004) and (3) they only considered enrolment and did not make any notice of success in higher education (Fine & Davis, 2003; Ou & Reynolds, 2010). In this chapter, we choose for the same approach as in the previous chapter by using a propensity score matching approach. Here, we will

calculate a propensity score for both early grade retainees (Grade 7 – Grade 9) and late grade retainees (Grade 10 – Grade 12), and we will consider the effect of both timings of retention on participation and short/long term success in higher education.

DROPOUT AND GRADE RETENTION IN FLANDERS: A FIRST LOOK AT THE DATA

In this dissertation, we only focus on grade retention and early school leaving in Flanders, Belgium. In the following paragraph, we further define some concepts and give an overview of the most recent statistics of grade retention and early school leaving in Flanders, based on administrative data. It is important to note that these data are not the same as the database we will use in the further analyses, but it gives an idea of the current situation. Moreover, the most recent dropout- and retention rate can act as a reference point when reading the different chapters.

First, it is necessary to define 'early school leaving'. Although this may seem unproblematic, there are different kinds of qualifications which should be taken into account. Generally, there are three different ways of defining early school leaving: (1) an early school leaver is a person that did not obtain a diploma of secondary education, (2) an early school leaver is a person that is not in compulsory education and did not obtain a diploma of secondary education in the academic, technical or artistic track, certificate of the 6th year of vocational education, certificate of part-time vocational education or apprenticeship, or a certificates of special needs education[iii] (a definition used by Van Landeghem, De Fraine, Gielen, & Van Damme, 2013) and (3) early school leavers are people between 18-24 who are not in education or training and who finished at most a lower secondary education. Throughout this dissertation, we will use the first definition. Although this definition is the most straightforward to implement in the analyses, we are aware of the drawbacks of this approach. One of the potential drawbacks lies in the fact that with this definition, the group of early school leavers can become very large, because, for example, students finishing the 6th year of vocational education do not get a 'diploma'. Compared to the second definition

(and to some extent, the third definition), the group of diplomas considered as a qualification is more limited, resulting in more early school leavers. Therefore, we will also consider the certificate of the 6th year of vocational education as a 'full' qualification.

The indicator of early school leaving with the highest precision, is the indicator based on the second definition. The high precision is due to the fact that this indicator is based on administrative data, and does not rely on survey data (with the corresponding sampling error and confidence intervals). With this definition, the dropout rate in Flanders in 2010 was at 13.9%. A closer look at these data already indicates some trends that we touched in the previous literature review. First, there is already a gender difference in the 'raw' percentages of early school leavers, with 11% of girls who left school before obtaining a qualification, compared to 16.6% boys (Van Landeghem, De Fraine, Gielen, & Van Damme, 2013). Second, older students seem to dropout more often compared to younger students (Van Landeghem & Van Damme, 2011a). Third, students with a different mother tongue than Dutch and/or students with a low-educated mother have a higher dropout rate compared to native Flemish students with a high-educated mother (Lamote, Van Landeghem, Blommaert, Nicaise, De Fraine, & Van Damme, 2013).

Concerning the rate of grade retainees, we have two indicators at our disposal: an indicator derived from the PISA-data and a Flemish indicator based on administrative data. In the PISA research project, every 15-year-old student had to answer the question "Have you ever repeated a grade?" and had to indicate in which grade. Based on this PISA-survey, Flanders has a rate of grade retention in lower secondary education (8.3%), 2% below the EU27 average (10.4%). Based on these data, we can only draw conclusion on the lower secondary education level. To draw conclusions on the higher secondary education level, and, more interestingly, say something about the rate of grade retainees in every grade, we can rely on the administrative data of the Flemish Ministry of Education. Based on this administrative data, we find that grade retention peaks in the 5th year of secondary education (8.76% of grade retainees). Overall, Flanders has 5.32% of grade retainees in secondary school, in the school year 2011-2012. Again, there are differences between native Flemish (majority) and non-native Flemish (minority)

students. The grade in which the minority group seems to struggle, is the 3th grade of secondary education, with a grade retention rate of 13%. Overall, in the school year 2011-2012, 8.9% of minority students repeated a grade in secondary school.

THE LOSO-STUDY

Although the data we gave in the previous section are very accurate, it is not possible to use them for further analysis. Administrative data are highly recommended for monitoring a certain phenomenon, but most of the time, not useful for thorough analysis because of the lack of background variables. These background variables are necessary to understand the complexity of a problem, and to draw conclusions about relations between student characteristics. Therefore, we turn away from these administrative data and throughout this dissertation, we will use data stemming from the longitudinal Flemish 'LOSO'-project (the Dutch acronym for Longitudinal Research in Secondary Education (Van Damme, De Fraine, Van Landeghem, Opdenakker, & Onghena, 2002)). This longitudinal research project started in 1990 and followed a cohort of 6,411 students in 90 secondary schools. These students were followed through secondary education, with a follow-up in higher education or in their first years on the labour market. The LOSO-dataset is characterized by its richness, containing data at the student, class, teacher and school level. Depending on the research questions, several variables were selected. Here, we discuss the main dependent variables, but more information (also on the independent variables) can be found in the respective chapters.

Early school leaving

To define early school leaving, we rely on the first of the three definitions formulated in the previous section. Consequently, in Chapter 2 and Chapter 3, an early school leaver was a student that left full-time education without a diploma of secondary education or a certificate of the 6th year of vocational education. This means that students changing over to part-time education were also considered as early school leavers. In our opinion, labelling these part-time students as early school leavers was a sound decision because part-time education is not an attractive alternative and students transferring to this part-time education often do so because

they did not yet reach the legal age to leave education (= 18 years). Although students in part-time education are able to obtain a diploma of secondary education, the number of students who succeeds in obtaining this diploma is negligible. When they obtain another certificate, this certificate does not guarantee the same outcomes on the labour market (Creten, Van de Velde, Van Damme, & Verhaest, 2004), and a lot of these students are still unemployed after 1 year (VDAB, 2013).

Grade retention

In Chapter 4 and Chapter 5, the operationalization of grade retention is simple and straightforward: when a student is in Grade X in Year Y, and the same student is in Year Y+1 still in Grade X, (s)he is a grade retainee. Because we have information on the position of (almost) every student for every year, we are able to define grade retention for every grade-level.

Engagement

In Chapter 3, we use engagement as a predictor for early school leaving. In line with the recommendations of Fredricks et al. (2004), we consider engagement as a multidimensional construct and we use a scale for behavioural engagement and for emotional engagement separately. Because there were no validated engagement scales in LOSO, we had to look for good indicators amongst the available scales. In search for a good scale, we used the definition and measurement suggestions of Fredricks et al. (2004), and relied on previous LOSO-studies that used an engagement measurement (Van de gaer, Pustjens, Van Damme, & De Munter, 2009). For behavioural engagement, we used a scale that measures 'attitudes towards homework'. This is in line with Fredricks' suggestion of measuring behavioural engagement: a behavioural engagement indicator should measure the conduct, work involvement or participation into school related activities and academic tasks (Fredricks et al., 2004). As an indicator for emotional engagement, we used a scale measuring the 'relationship with teachers'; a common conceptualization of emotional engagement. We did not find a good indicator for cognitive engagement. For the specific items of each scale, see Appendix 1. Both scales stem from a well-being questionnaire, which was administered at four different moments: at the end of the first, second, fourth and sixth year of secondary

education (for students retained in grade, one year later at the end of the second, fourth and sixth year).

Language achievement

In Chapter 4, the Dutch language achievement was one of the main dependent variables. This language achievement was tested five times throughout a student's career in secondary education: at the beginning of the first year, and at the end of the first, second, fourth and sixth year (and again three times for the delayed students). Although different versions of curriculum-relevant tests were administered in the different grades and tracks (with adapted difficulty levels), all tests shared overlapping items, allowing for the construction of a common scale by using Item Response Theory (IRT). Hence, this scale is used for both cross-sectional analysis (between tracks), as well as for longitudinal analysis (growth). In Chapter 4, we used four different measurement occasions (excluding the measure at the beginning of the first year) for the estimation of growth curves.

Academic self-concept

The second dependent variable used in Chapter 4, is the academic self-concept of the student. With this variable, we try to grasp the general academic self-concept, instead of domain specific self-concept measures which are usually perceived as superior (but not available in LOSO; for a thorough discussion on this issue, see: Pinxten, De Fraine, Van Damme, & Van Den Noortgate, 2013). This scale stems from a general well-being questionnaire which was administered four times, together with the achievement tests (but not at the beginning of the first year).

Post-secondary education outcomes

While the majority of dependent variables throughout this dissertation was measured at secondary school level, we focus in Chapter 5 on outcomes at post-secondary education level. At this level, we consider three different outcomes: participation in higher education, and success in the short- and the long-term. Determining whether a student participates in higher education was straightforward: when a student was subscribed in any form of higher education (professional or academic) at the beginning of an academic year, (s)he participated in higher

education. The success of a student in the short-term was evaluated by looking at the results of a student's exams (and potential re-examination) in the first year. When a student passed *all* the exams and was allowed to follow the study program of the second year, this was considered as short-term success. To evaluate success in the long-term, we followed the same reasoning, but we focussed on the outcome at the end of the third year in higher education. We opted for the third year, because at the end of this year, in the actual context, a diploma can be awarded (Professional – or Academic Bachelor). Consequently, a student was considered as successful in the long-term when he received a Bachelor's degree[iv] after 3 years.

While the operationalization of participation in higher education is univocal, the operationalization of 'success' is debatable, for two reasons. First, the success-variable was constructed as a dichotomous variable: successful vs. not successful. In reality, it can also be that a student passes some exams and fails for others. If that is the case, the student can already enrol in some courses of the subsequent year. However, in LOSO, this kind of information was not available, so we had to (over?)simplify the idea of 'success'. A second problem with the 'success'-variable in Chapter 5, relates to the timeframe that we considered. If a student makes a wrong choice of study program in his first year, and changes to another study program in the second year, he will (most likely) have one year of delay and cannot be successful anymore based on our operationalization. Unfortunately, we only have (reliable) data on the success of a student for 3 years after graduation from secondary education. Therefore, the success-variable should be interpreted with some caution.

METHODOLOGY

Throughout this dissertation, we use state-of-the-art analyses in order to draw reliable conclusions. First, we always take the multilevel nature into account. In the first two papers, we use a (discrete-time) survival analysis, while in the third and fourth paper, we combine a propensity score analysis with a growth curve and/or regression analysis. In this introduction section, we give a short description of each

method, but more information on these methodologies can be found in the respective chapters.

Multilevel models

The data used in the following chapters, had a typical hierarchical or clustered structure with measurements nested within students (Chapter 4) and, classic in educational research, students nested within schools (Chapter 2, 3, 4, 5). This grouping of students is not ignorable and ignoring it, for instance by conducting a one-level (linear) regression, would violate the assumption of data independence of the residuals. As Goldstein describes it: the group [school] and its members [students] both influence and are influenced by the group membership (Goldstein, 2011, p. 2). The proper way of handling such data structures, is by making use of multilevel models. Using a multilevel model has several advantages. To name just two of them: it is possible to decompose the variance across the different levels, and it gives a better estimate of the standard errors of the regression coefficient estimates, at least if every level is taken into account.

Throughout the different chapters, we account for the hierarchical structure in different ways: by using multiple membership and cross-classified models in Chapter 2, by the use of a so-called 'sandwich-estimator' (Muthén et al., 2002) in Chapter 3, by using multilevel growth models in Chapter 4 and by using multilevel logistic regressions in Chapter 5.

Survival analysis

In Chapter 2 and Chapter 3, we were not only interested in 'whether' a student drops out, but also 'when' this student drops out. Following Singer and Willett (2003), we answer this question with a survival analysis, and to be more specific we use a discrete-time survival analysis. With this survival analysis, we are able to estimate the probability to drop out for every grade separately. However, to estimate this probability accurately, we use a discrete-time survival analysis instead of treating time as being continuous. We use this discrete-time approach because a lot of students drop out at the end of a school year, and estimating this in a continuous framework would create the problem of tied observations, which can

lead to biased estimates (Allison, 1982). In the subsequent chapters, we model this discrete-time survival in two different frameworks: in a logistic regression framework and in a latent class regression framework.

Matching

In Chapter 4 and Chapter 5, we are interested in the effect of an intervention on several outcomes. In answering such kind of questions, a randomized controlled trial (RTC) is still considered as the 'golden standard', because treatment assignment is done randomly and is not influenced by background characteristics (measured or unmeasured) of the subject. In an RTC, the effect of a treatment is simply the difference in outcome between the treated and the control group. In Chapter 4 and 5, we estimate the effect of a popular educational 'treatment', grade retention, on achievement, academic self-concept and post-secondary education outcomes. For a good evaluation of this treatment (grade retention), an experiment in which we randomly assign students to a treatment or control group, meaning that we – without considering background characteristics – randomly decide which student has to repeat a grade and which student is promoted to a higher grade, is necessary. Of course, this is ethically not recommended. As a consequence, we cannot evaluate the effect of our treatment by simply comparing the average outcomes of the retained group with the outcomes of the promoted group, because students in both groups can have very different pre-treatment background characteristics. A solution to this problem, is to match a student of the treated group with a student of the control group with comparable pre-treatment background characteristics. This is feasible when the number of variables to match on is limited, but once this number increases, this matching becomes nearly impossible. In that case, a balancing score can offer the solution. One of the most popular balancing scores, is the propensity score. This propensity score is the "(…) conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum & Rubin, 1983, p. 41). The calculation and implementation of this propensity score follows a stepwise procedure. First, we select the relevant covariates (i.e. covariates that are (theoretically) related to the treatment and the outcome) for calculating the propensity score (Rubin, 2008). Second, with the selected covariates, we calculate the propensity score for each

student. We do this by regressing the covariates on the treatment status, so that the propensity score is the predicted probability of the treatment based on the regression model. The third step is the actual matching: we match students in the control group with students in the treated group with a similar probability of treatment. These matched students are then used in subsequent analyses. In Chapter 4 and 5, we provide more details on the matching algorithms, the selection of variables and quality issues of the matchings.

Chapter 2

Dropout in secondary education: an application of a multilevel discrete-time hazard model accounting for school changes

Published as: Lamote, C., Van Damme, J., Van Den Noortgate, W., Speybroeck, S., Boonen, T., de Bilde, J. (2013). Dropout in secondary education: An application of a multilevel discrete-time hazard model accounting for school changes. *Quality and Quantity*, *47* (5), 2425-2446.

ABSTRACT

For several decades, researchers have focused on dropout in search for an explanation and prevention of this phenomenon. However, past research is characterized by methodological shortcomings. Most of this research was conducted without considering the hierarchical structure of educational data and ignored the longitudinal path towards dropout. Moreover, research that did take into account these shortcomings, did not correct for student mobility between schools, despite the strong correlation with dropout (South, Haynie, Bose, 2007). In this study, we attempt to address these shortcoming by implementing a multilevel discrete-time hazard model and exploring the effect of different school classifications on the school effects. Partially analogous to Grady and Beretvas (2010) we compare models with estimated school effects based on the first and on the last school attended and compare these models with multiple membership models and cross-classified models. The results of this comparison indicate that ignoring student mobility can have strong implications on the predictors of dropout. Not only do models which take into account this mobility yield better model fits, models ignoring this mobility tend to miss the effect of school level variables. With respect to the conclusions on dropout research, our models provide evidence for the often cited student characteristics predicting dropout and indicate stronger school effects than generally assumed.

INTRODUCTION

For the majority of students, graduating from secondary education can be seen as obvious and unproblematic, at least in industrialized countries. However, a part of the students never reaches this endpoint and leaves secondary education before receiving a diploma. These students are termed as dropouts. Over the past 25 years, numerous studies have focused on the dropout phenomenon in secondary and also in higher education (e.g. Finn, 1989; Lee & Burkam, 2003; Rumberger, 2001; Roderick, 1994). The focus on dropout in secondary education can be explained by the severe consequences of this dropout: in an age where diplomas continue to gain in importance, leaving school without a qualification implies a strong decrease in future employment opportunities (Solga, 2002). To classify the numerous variables considered in past research on dropout, Rumberger (2001) distinguished two broad perspectives to look at student dropout: an individual and an institutional perspective.

The individual perspective encompasses the majority of regularly observed explanatory variables for dropout. Several of these variables are also used in this study and include gender, (prior) achievement, grade retention, student mobility and student engagement. One of the most commonly observed student characteristics predicting dropout, is gender. Although regularly observed, the effect of gender on dropout was not always unequivocal: a first set of studies found that boys tend to have higher odds of leaving school (e.g. Alexander et al., 2001; Marks, 2007), a second set concluded an opposite effect with girls having higher odds of leaving school (Croninger & Lee, 2001; Goldschmidt & Wang, 1999) and a third set of studies found no effect of gender at all (Lee & Burkham, 2003; Hickman et al., 2008). The strong relation between dropout and grade retention in primary and secondary school received also a lot of attention (e.g. Alexander et al. 2001). In their review, Jimerson et al. (2002) pointed on the strong predictive power of (early) grade retention on dropout. For example: Rumberger (1995) concluded that students who repeated a grade, were nearly 11 times more likely to drop out compared to students who had not repeated any grade. Besides gender and grade retention, low-achievers are often at higher risk for dropout. Alexander et al. (2001) concluded that low achievement in first grade, was one of the major predictors of

dropping out. Another relevant predictor of future dropout, is student mobility. Students who experienced one or more school changes, are approximately twice as likely to drop out compared to non-mobile students (South et al., 2007). Dropping out of secondary education is not only a matter of gender , grade retention, or low school performance, but can be an expression of a more subtle and underlying process of reduced school engagement. School engagement was included in many dropout theories, but not often considered in empirical research (see e.g. Ream & Rumberger, 2008). However, Finn's (1989) identification-participation model stated that active participation in school and a feeling of identification with school, strongly mediates school outcomes.

In addition to the individual perspective, also the institutional setting in which students live can predict future dropout. Regarding this institutional perspective, we consider the family background of the student and the schools attended. The family background is usually operationalized by the socioeconomic status (SES) of the student. As for SES, especially the educational level of (one of) the parents and their income remain an important contributor to success in school. Even when controlling for other student characteristics, students with a low SES background have a higher dropout probability (Alexander et al. 2001). Concerning the second institutional variable, research is much more limited as only a few studies included school characteristics in a multilevel model, which led to several new findings (Goldschmidt & Wang, 1999; Lee & Burkam, 2003; Luyten et al., 2003; Rumberger 1995). Lee and Burkam (2003) focused on school academic and social organization and concluded that students were less likely to drop out in schools where relationships between teachers and students are more positive. Furthermore, Rumberger and Thomas (2000) concluded that high-SES schools had 40% lower dropout rates than average SES schools, whereas low-SES schools had 60% higher dropout rates than an average SES school.

Rumberger (2001) stated that dropout is a final action in a long-term process. Nevertheless, several of the mentioned studies have a narrow focus on the dropout phenomenon itself (e.g. Luyten et al. 2003). Indeed, research should focus on the longitudinal nature of dropout, instead of drawing conclusions based on one moment, since not every predictor seems to have the same effect at every moment.

Only few studies took this longitudinal nature into account. Bowers (2010) concluded that time-varying variables in a longitudinal study (e.g., grade retention, achievement scores) predicted dropout more than time-invariant variables (e.g., gender).

Besides ignoring the longitudinal nature of the data, research also often ignores several aspects at the institutional or school level. Such school characteristics, as well as student characteristics, can be integrated into a hierarchical model, with students nested in schools. As mentioned earlier, this multilevel approach was rarely used.

To account for both the longitudinal and hierarchical structure of the data, we propose using a longitudinal multilevel model. A complicating factor however, is that a pure nesting of students within schools is not always in accordance with educational reality, where students encounter school changes during their career through (secondary) education. When focusing on school effects, changing schools entails a change of school environment and thereby other values for the school level variables. Students who attended multiple secondary schools, have been exposed to multiple school effects (Grady & Beretvas, 2010), and most of the time these school changers are no random group and should therefore not be excluded from a dataset (Goldstein, 1997). As mentioned earlier, drop-out was found to be related with student mobility, pointing to the importance of including mobile students in a study on dropout. This change extends the usual nested structure of educational data, and suggests the use of cross-classifications and multiple membership models (Goldstein, 2011).

RESEARCH QUESTIONS

Taking into account the aforementioned methodological considerations, and the theoretical background of dropout in secondary education, we propose four research questions. The first research question is the main interest of this study and concerns the methodological approach in studying dropout, where we are interested in the most appropriate model for modelling dropout in secondary education. More

specifically, the first research question deals with the longitudinal character of the data and the multiple schools attended. We will compare a model with a multiple membership structure, a model with a cross-classification and models only focusing on one school. We expect to find a better model fit when taking all the schools attended into account. We will also compare the methods regarding the answers they give on the three following substantive research questions. In the second research question, the effect of student characteristics on the chance to dropout is addressed. In line with past research on dropout, we expect a higher chance to dropout with male students, who repeated a grade in primary and/or secondary education and grew up in a low SES family. The third research question addresses the effect of school level variables on the chance to dropout. Although previous research rarely focused on school level variables, the research that did, found, among others, effects of school composition, school size and relationships. The fourth research question follows directly from the longitudinal approach of modelling dropout in this study and will identify the more hazardous grades for dropout in secondary education.

The choice of the variables on the student level and school level is mainly based on the research literature on dropout, and therefore the study will partially replicate previous studies. The main contribution of this study is, as mentioned, the implementation of alternative models accounting for the longitudinal character, the multilevel structure and the school changes. The next part presents the dataset, which we used to test the proposed models and from which we will draw our overall conclusions. The description of the dataset is followed by a short introduction into hazard modelling, multiple membership and cross-classified models in the method section. We end the method section with an integration of these multiple membership and cross-classified models into a hazard analysis.

DATA SOURCE

The data were drawn from the Flemish 'LOSO'-project (Van Damme et al., 2002). This longitudinal research project started in 1990 and followed a cohort of 6,411 students in 57 secondary schools. These students were followed through secondary

education, with a follow-up through higher education or the first years on the labour market. The LOSO-dataset contains data at the student, class, teacher and school level. In this study, we selected a subsample of 4,735 pupils in 55 secondary schools, of which 514 dropouts, who started in 1990 in the first year of (regular) secondary education (grade 7) and had known educational positions and school ID's for every year through secondary education. Students who repeated a grade and/or changed schools were also included. Because of our interest in the effect of school characteristics on dropout rates, and because we have information on school characteristics for LOSO-schools only, we excluded students transferring to a school that was not included in the LOSO-project. The LOSO-project selected secondary schools from several regions in the Flemish community, and in the selected regions, almost every school participated. Because of the strong coverage within every region, we were able to follow-up the students who changed schools within the region and we can assign school characteristics to every student for every time point.

Measures

A student is considered as dropped out at the moment (s)he leaves fulltime secondary education, and is not able anymore to obtain a diploma of secondary education (a full high school diploma). This definition of dropout is very strict and as a result, students who leave fulltime secondary education for part-time education or for other alternative schoolings are considered as dropouts although in the Flemish community they still can receive a qualification. This qualification however is not comparable to a diploma of secondary education, since such qualification does not guarantee the same chances in higher education and on the labour market.

We selected five student background characteristics that will be included in our models as independent variables: gender (GEN), initial cognitive ability (COGN), socioeconomic status (SES), older at start of secondary education than the norm student (RET_P) and retained in grade during secondary education (RET_S). Gender (GEN) is implemented as dummy variable, with value '1' for girls, and boys with value '0' as reference category. The initial cognitive ability (COGN)

includes three components: math achievement, language achievement and intelligence, measured at the start of grade 7. All the components had high internal consistencies, ranging between $\alpha = .82$ and $\alpha = .93$ (Van Damme et al. 1997). The math achievement test was a multiple-choice test which consisted of 50 items, assessing set theory, algebraic problems and geometry. The raw scores were converted into item response theory (IRT) scores. The content was in line with the content received in the highest grades of primary school. The language achievement test consisted of 100 items, divided into 4 sections: spelling, grammar, language use and comprehension. The intelligence score was based on the Getlov-intelligence test, which measured the intelligence of students at the beginning of secondary education. This test comprises three test scores: a score for verbal intelligence, a score for spatial representations, and a score for numerical intelligence. A factor analysis on the 5 intelligence and achievement scores, gave a one factor solution. The initial cognitive ability predictor we will use (COGN) are factor scores resulting from this factor analysis. Students with missing values on one or more of the 5 scores were excluded from further analysis. The socioeconomic status (SES) is based on parental education, parental occupational status, monthly income and a variable measuring the cultural capital of the family. These variables were combined into a factor analysis, resulting in a one factor solution. Because we only have self-reported information on grade retention in primary education, we deduct whether or not the student was retained in grade from his age (RET_P). Traditionally, students start in secondary education in the year they reach the age of 12 years. Students who were older at the start of secondary education, most likely repeated a grade in primary education and are coded '1'. Students who started secondary education at the age of 12 (or younger), are coded '0'. Grade retention in secondary education (RET_S) is implemented as a time-varying dummy variable, with value '1' in the grade in secondary education which the student repeated and in the subsequent grades.

Besides the student background characteristics, we included the following school composition variables: socioeconomic school composition (SES-S), cognitive composition of the school (COGN-S) and gender composition of the school (GEN-S). These variables are aggregations of the student-level variables

SES, COGN and GEN. Although we only incorporate students belonging to the LOSO-cohort in this study, the construction of these aggregated variables is based on all the students for which this information was available, including non-LOSO students, as recommended by Snijders and Bosker (1999).

An indicator of the school climate, referring to the relationship with teaching staff and peers, reported by the students (RELAT-S) was also included. As Fredricks et al. (2004) mentioned, the relationship between student and teachers on the one hand, and the relationship between student and peers on the other, is a strong indicator of the emotional engagement towards the school. Therefore, we implemented this RELAT-S. This variable is part of a larger set of school climate variables, constructed by De Fraine (2003), and was also implemented in Opdenakker and Van Damme (2005). RELAT-S consists of three first order variables: (1) relationship with teachers, with items as "Teachers trust the students" and "Teachers give personal attention to every student"; (2) granting equal rights to every student, with items as "Teachers behave differently depending on the track" and (3) friendships amongst students, with items as "New students are accepted quickly". Because of the strong correlation between these three variables, they were combined into a new, second order variable relationships at school.

METHOD

In this study, we are not only interested in 'whether' students drop out of secondary education, but also in 'when' students drop out. As Singer and Willett (2003) suggested, this question is answered by making use of a discrete-time hazard analysis (also known as discrete-time event-history analysis or discrete-time survival analysis). However, making use of standard discrete-time hazard techniques would ignore the fact that students are clustered in schools and tend to change schools during secondary education.

In the next sections, we discuss the multilevel discrete-time hazard (MDTH) approach and extend this approach by taking into account student mobility, making use of a multiple membership model and models with cross-

classifications. First, we discuss the benefits of this MDTH approach compared to the more 'traditional' approach in studying dropout, followed by a short introduction to multiple membership models and models with cross-classifications. We end this method section with the integration of the multiple membership structure and the cross-classifications into the MDTH model.

Multilevel discrete-time hazard model

In this study, we opt for a multilevel discrete-time hazard analysis (MDTH) for four main reasons: the interest in predicting the (non) occurrence of an event (here: dropout), the longitudinal character of our data, the problem of tied observations and the hierarchical structure of educational data.

Hazard analysis is a flexible (and straightforward to implement) tool which measures the (non) occurrence of an event, as well as the timing of the event. This time-to-event analysis starts at a well-defined starting point (here: at the beginning of grade 7) and ends at the occurrence of the targeted event (here: dropout), predicting the event-occurrence at each time-interval. However, because of the longitudinal character of our data we may encounter a common aspect of longitudinal data analysis: censoring. Censoring occurs when some participants do not experience the studied event during the time period of the study. In educational context, when studying dropout, this censoring might occur in different manners: students did not experience the event when the study ended, students were lost to follow-up or students withdrew from the study because of another event not of interest (e.g. death) (Kleinbaum & Klein, 2005). In our study, censoring can only occur at Grade 12, when students did not experience the targeted event and stayed in the dataset until graduation. For the hazard analysis in this study, a discrete-time approach seems the most appropriate method, since our data encloses multiple grades. Most of the individuals have a dropout date that matches the end of a school-year, or the beginning of a subsequent school-year, which raises the problem of tied observations (two or more students experience the event at the same time). Such tied observations are not uncommon in educational research, where events are often observed or take place at the end of a school-year. Using a continuous-time approach despite tied observations can lead to biased estimates (Allison, 1982). A

possible and easy to implement solution is to treat time as if it were discrete. Although not every student drops out at the last day of the school-year, this discrete-time approach seems a good approximation of the moment of dropout. The discrete-time hazard analysis allows us to use the logistic regression model, with the addition of time-dummy variables for every grade and time-invariant and/or time-varying covariates.

Traditional discrete-time hazard analysis assumes that students behave independently from each other. This is not the case in an educational context, where students are grouped in classes and schools, and observations of students within the same school are not independent from each other. Therefore, a multilevel structure with individuals at Level-1, grouped into Level-2 units (schools) is more appropriate. As such, the multilevel discrete-time hazard model is written as (Goldstein, 2011; Barber, Murphy Axinn, & Maples, 2000):

$$\log\left(h_{tij}/\left(1-h_{tij}\right)\right) = \beta_0 + \alpha_t + \beta x_{tij} + u_j \qquad (1)$$

where $h_{tij} = Pr\left(Y_{tij} = 1 \middle| Y_{t-1,ij} = 0\right)$ is the hazard that student i in school j drops out at time t (given the event has not yet occurred to that individual before time t), α_t is a function of time, β a vector of parameters representing effects of covariates x on the probability of event occurrence and u_j the school-level random effect.

In a traditional multilevel discrete-time hazard model, students are clustered in only one school and consequently, researchers need to allocate only one school ID for every pupil. This allocation is straightforward for students attending only one school during the study, but becomes problematic for students changing schools. For these school changers, several ad-hoc approaches for the choice of school ID's have been adopted in previous research. One of the possible solutions is to take the school ID of the last or the first school attended. Using the school ID (and corresponding school characteristics) of the first or last school, however, substantially underestimates the true between-school variance (Goldstein, 2011; Goldstein, Brugess, & McConnell, 2007; Grady & Beretvas, 2010). Another

solution for this school change is to simply delete the mobile students. The resulting reduced dataset, however, may be biased as most of the time, this group of school changers is no random group (Goldstein, 1997).

The following sections go more deeply into the matter of modelling student mobility, by making use of multiple membership models and cross-classified models. The resulting data structure is not strictly hierarchical anymore, but non-hierarchical in which repeated measures are cross-classified by students and schools. Although the benefits of these approaches are widely studied (see e.g. Grady & Beretvas, 2010; Luo & Kwok, 2009; Fielding, 2002), only few educational studies took this non-hierarchical structure into account (e.g. Goldstein & Sammons, 1997; Pustjens, Van de gaer, Van Damme, & Onghena, 2008; Raudenbush, 1993; Teitler & Weiss, 2000). In this study, we give a short introduction into these approaches, implement these into a discrete-time hazard analysis and present an application of these approaches.

Multiple membership models

Because many students change school at least one time, we could incorporate this school movement by making use of a non-hierarchical approach where the lower level units (here: students) are members of more than one higher level unit, resulting in a multiple membership model.

Because we have a school ID for every school in every grade, and we know the grade of dropout, we are able to assign weights to every secondary school the student attended. By assigning weights proportionally to the time spent in every school, we implicitly assume that each year the school has the same importance. One could argue that the last school attended has a more substantial impact compared to the first school attended, and should therefore receive a higher weight, irrespective of the time spent in the last school. On the other hand, Browne, Goldstein, and Rasbash (2001) and Goldstein et al. (2007) experimented with different weighting schemes, and concluded that the parameter estimates are (relatively) insensitive to the choice of weighting scheme.

Taking into account this multiple-membership, the basic hierarchical logistic model is extended to a model with a multiple membership structure (Goldstein, 2011):

$$\text{logit}(\pi_{ij}) = \beta_0 + \sum_{j \in school(i)} w_{i,j} u_j \tag{2}$$

$$\text{with } u_j \sim N(0, \sigma_u^2) \text{ and } \sum_{j \in school(i)} w_{i,j} = 1$$

In Equation (2), $w_{i,j}$ is the weight assigned to the random effect for student i in school j. For each student, the sum of these weights equals 1. As a result of these weights, we must be careful when interpreting the variance components of the effects. As Leckie (2009) states, the variance σ_u^2 of the school effect varies along with the mobility of the student. A student who attended only one school, has a contribution of the school to the variance of the student of σ_u^2, but students who experienced school change will have a contribution of the school to the variance of the student, proportional to the time spent in a particular school. Consider for example a pupil who spends 4 years in school A and 2 years in school B, the model is as follows (Goldstein, 2011):

$$\text{logit}(\pi_{ij}) = \beta_0 + w_{1i,jA} u_{jA} + w_{2i,jB} u_{jB} \tag{3}$$

$$\text{with } w_{1i,jA} = 4/6 \text{ and } w_{2i,jB} = 2/6$$

and in calculating the total contribution of the attended schools, the weighted effects of every school attended are summed to:

$$\sigma_u^2 = w_{1i,jA}^2 \sigma_{uA}^2 + w_{2i,jB}^2 \sigma_{uB}^2 \tag{4}$$

This is the strength of these multiple membership models: the longer a student spends in one school, the greater the impact of that particular school in the final school effect.

When these models are extended with variables on the school-level, this multiple membership must also be taken into account. However, computing school-effects with a multiple membership is, with the current software limitations, not obvious and time-consuming. One has to weigh each variable on the school-level with the corresponding weightings for the attended schools, for a particular student. After this weighing, the variable on the school level equals the sum of the weighted school-characteristics for every school attended. For example: a student who spent 2 year in a school with 90% girls and 4 years in a school with 69% girls will have an outcome on 'school composition by gender' of $(2/6) \times (0.9) + (4/6) \times (0.69) = 0.76$ (De Fraine, 2003).

Cross-classification models

A drawback of the multiple memberships are the practical difficulties. Moreover, multiple membership models give information on the time spent in a particular school, but do not give any information on the moment the student was in a particular school. To model these school changes, we can also use a cross-classified model, where the school attended is not a fixed feature of the student, but depends on the grade attended by that student. This means a reformulation of the traditional multilevel discrete-time hazard model with students at Level-1 and schools at Level-2 to a model with observations (Level-1) nested in a combination of students and schools at Level-2. This way, schools are not characteristics of the students at Level-2, but are a characteristic of the Level-1 unit defined as the measurement moment for a student. Such a formulation reflects the model described by Raudenbush and Bryk (2002) and Fielding and Goldstein (2006) in which repeated measures are cross-classified by students and schools.

With this cross-classification, there is no need to weigh every school characteristic because a difference or change in school characteristic due to school change is directly captured. The basic logistic cross-classification model is written as (Goldstein, 2011):

$$\text{logit}(\pi_{ij}) = \beta_0 + u_{1i} + u_{2j} \tag{5}$$

$$u_{1i} \sim N(0, \sigma^2_{u_1}) \qquad u_{2j} \sim N(0, \sigma^2_{u_2})$$

Equation (5) is the basic cross-classified model with $\sigma^2_{u_1}$ as the variance between students and $\sigma^2_{u_2}$ the variance between schools.

Multiple Membership and Cross-classified Discrete-time Hazard Model

The aforementioned methodological approaches provide a strong framework for analysing dropout in secondary education. As mentioned, changing schools is a strong predictor of this dropout and should be implemented in a model which makes use of the hazard approach to model dropout. When we incorporate the multiple membership in the discrete-time hazard model, the model will be:

$$\log(h_{tij}/(1 - h_{tij})) = \beta_0 + \alpha_t + \beta x_{tij} + \sum_{j \in school(i)} w_{i,j} u_j \tag{6}$$

with α_t as a function of time, βx_{tij} as a vector of parameters representing effects of covariates on the probability of event occurrence and with the multiple membership part, represented by summation of the weighted school characteristics.

When we incorporate the cross-classification in the discrete-time hazard model, the model will be:

$$\log(h_{tij}/(1 - h_{tij})) = \beta_0 + \alpha_t + \beta x_{tij} + u_j \tag{7}$$

Indeed, this cross-classified model is written in the same way as the traditional MDTH with only a random school effect and no random student effect, because a student can drop out only once and variance between students is already accounted for by the first level Bernouilli distribution. The cross-classified aspect is in the handling of the nesting of students within schools. In the traditional MDTH, a student is nested in only one school, where in this cross-classified approach, a measurement is nested within a crossing of a student and a school, and as such this school indicator and corresponding school characteristics can change over time. The latter is the main raison for adopting a cross-classification; the school characteristics are handled in a similar way as in a 'traditional' cross-classification.

We will estimate different models, each model with a different approach for the school ID, partially analogous to Grady and Beretvas (2010). The different approaches are respectively the first school approach (by taking the school ID of the first school attended), the last school approach (by taking the school ID of the last school attended), the delete approach (by deleting mobile students), the multiple membership and cross-classified approach. We choose these different approaches because of the frequent implementation in previous research and to illustrate the possible bias of these approaches. These models are preceded by a model without taking into account the school level. All the models were estimated with the MLwiN software (Rasbash, Steele, Browne, & Goldstein, 2009), with the Markov Chain Monte Carlo estimation. This MCMC-approach can easily handle the multiple membership and cross-classification. In order to compare model fits, we use this MCMC-approach also for the preceding models. The evaluation of model convergence was based on the trajectories of the estimates, the autocorrelation function and the Raftery-Lewis diagnostic as a diagnostic for the required length of the Markov chain. We compare the different model fits by making use of the deviance information criterion (DIC), where models with lower DIC values are preferred. Differences around the value of 10 or higher, are seen as substantial (Spiegelhalter, Best, Carlin, & van der Linde, 2002).

In the next part, we will present different results, arising from different modelling approaches.

RESULTS

Hazard analysis of dropout

As one of the main features of hazard analysis, we first present a table of the number of students who drop out from the beginning of secondary education, up until the end of secondary education (Table 1), a table that in hazard analysis is called a life table. The columns of hazard and survival are essential for hazard analysis. This hazard h is the conditional probability that an individual i will experience the event of interest in a time period t, given that they did not experience the event in an earlier time period (Singer & Willett, 2003). In this study, the hazard probability is a representation of the proportion of dropouts in the group of students in a certain grade which are still at risk for dropout. The interpretation of this probability is straightforward: the higher the hazard probability in a particular grade, the higher the risk for dropout in that grade. The survival estimate on the other hand, does not represent a probability in a certain grade but represents the probability that a student will 'survive' (will not experience the event of dropout) until a certain point in time and is an accumulation of the preceding hazard estimates.

A closer inspection of these hazard and survival estimates gives information on the patterns of dropout in the different grades, controlling for students who already left secondary education. At grade 7 this hazard for dropout is rather low, but gradually, this hazard increases in grade 8 until grade 10 with a peak at grade 11 and a decline in grade 12. The survival probability indicates that 89% of the students in this dataset did not experience the event and graduate by the end of grade 12.

Table 1: Life table

Grade	Enrolled at beginning of year	Not in subsequent year	Censored	Hazard h	Survival s
0	0	0		0.0000	1.0000
7	4,735	2		0.0004	0.9996
8	4,733	63		0.0133	0.9863
9	4,670	111		0.0238	0.9628
10	4,559	135		0.0296	0.9343
11	4,424	143		0.0323	0.9041
12	4,281	59	4,222	0.0138	0.8917

School change prior to dropout

Changing schools during secondary education is substantial in Flanders. More than one quarter of the students in our dataset changes school once or more. This school change takes on an even larger proportion with students who drop out. Half of these dropouts changes school at least one time. An overview of school changes distributed over graduates and dropouts can be found in Table 2.

Table 2: Number of school changes

Number of school changes	Graduate		Dropout	
Never	3167	*75%*	259	*50.4%*
1 time	943	*22.4%*	206	*40.1%*
2 times	100	*2.4%*	45	*8.7%*
3 times	11	*0.2%*	4	*0.8%*
Total	4221	*100%*	514	*100%*

In the following paragraphs, we present the results on the research questions. At first, we compare different baseline hazard models. These baseline hazard models were estimated by making use of the first school approach, last school approach, delete approach, multiple membership approach and cross-classification approach. We compare these baseline hazard models on their DIC-values.

Baseline hazard models

In Table 3, the first set of models are baseline hazard models, with only time-variables (Grade) as predictor for dropout. These time-variables tend toward an increase per grade until grade 11 in the risk for dropping out, with a decrease in 12^{th} grade. In the first model, we ignored the fact that students were nested within schools and thereby (partially) replicated past research on dropout in secondary education (see e.g. Bowers, 2010; Roderick, 1994). From the second model on, we adjusted for the fact that every student is in a certain school for a certain time. As such, the second model follows the traditional first school approach, where random

effects are based on the first school the student attended. The DIC-values indicated an improvement of the model (Model 2 DIC = 4,572.02), compared to the first model (Model 1 DIC = 4,932.38), suggesting that there is variation between schools in the dropout-rate. In the third model, we took the last school attended as school classification. Comparing this with the second model, this last-school approach gives a better model fit (Model 3 DIC = 4,422.35), favouring this nesting above the first school attended.

In the fourth model, we explicitly modelled school change by making use of a multiple membership structure. Less expected on the basis of the literature, this multiple membership model returned a higher DIC-value, indicating a poorer fit (Model 4 DIC = 4,558.50). There are two possible explanations: first, only the last school attended matters. Secondly, not every school attended has the same impact on dropout, and the last schools have a greater impact compared to previous schools. This brings us to the discussion by Goldstein et al. (2007) of the choice of weighting scheme. Instead of modelling this multiple membership model with different weighting schemes, we changed over to a model where time is nested within a cross-classification of a student in a certain school. Based on the DIC-value, we prefer the model with cross-classifications above the first school or last school approach (Model 5 DIC = 4,375.43). In the last model in Table 3 we simply deleted students who changed schools. Because this deletion resulted in a reduction of 30% of the original dataset and nearly 50% of the dropouts, we cannot reliably compare the DIC-value of the delete approach with the previous models.

When we focus on the fixed effects parameter estimates, some differences arise. Compared to using a cross-classified model, ignoring the nested structure of students within schools, results in larger effects of the time-variables on dropout. This overestimation is reduced by incorporating a school-level which partitions a proportion of the variability to the school level. The variance estimates of the random effects differ much stronger. The first school approach (Model 2) resulted in a much smaller school level variance (Model 2 σ_u^2 = 1.109) compared to the approaches explicitly modelling the mobility. Following Goldstein (2011), the small(er) school effect in the first-school approach could be expected because almost 28% of the pupils changed school at least one time: the limited effect of the

first school on these mobile students results in a decrease of the average between school variance. When we ignore this mobility, as is the case in the first school approach, we assume that the first school attended contributes to the same extent for mobile students as for non-mobile students, which is clearly not the case. The last school approach (Model 3) yields a higher school level variance (Model 3 σ_u^2 = 2.199) compared to the multiple membership model (Model 4 σ_u^2 = 1.334), but a lower school effect than the cross-classified approach (Model 5 σ_u^2 = 2.263). A possible explanation for these differences lies in the moment of school change or school attendance. It seems the case that not only the time spent in a particular school is of importance, but also the timing and order the student was in different schools.

Table 3: Baseline Hazard models[v]

	Model 1 Without school level		Model 2 First school		Model 3 Last school		Model 4 Multiple Membership		Model 5 Cross-classified		Model 6 Delete approach	
	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Constant	-7.580	*0.619*	-7.779	*0.648*	-8.183	*0.693*	-7.921	*0.603*	-8.023	*0.669*	-9.234	*1.209*
Grade 8	3.221***	*0.632*	3.218***	*0.640*	3.248***	*0.665*	3.160***	*0.593*	3.160***	*0.640*	4.362***	*1.191*
Grade 9	3.796***	*0.627*	3.823***	*0.634*	3.853***	*0.659*	3.762***	*0.587*	3.668***	*0.634*	4.632***	*1.190*
Grade 10	4.031***	*0.625*	4.118***	*0.633*	4.153***	*0.658*	4.050***	*0.586*	3.936***	*0.633*	4.937***	*1.188*
Grade 11	4.124***	*0.625*	4.225***	*0.633*	4.291***	*0.658*	4.195***	*0.585*	4.058***	*0.633*	5.022***	*1.188*
Grade 12	3.312***	*0.633*	3.446***	*0.640*	3.508***	*0.665*	3.401***	*0.594*	3.257***	*0.641*	4.063***	*1.198*
Between schools σ_u^2			1.109	*0.290*	2.199	*0.663*	1.334	*0.363*	2.263	*0.637*	2.500	*0.793*
DIC	4,932.38		4,572.02		4,422.35		4,558.50		4,375.43		2,370.05	
N	28,585		28,585		28,585		28,585		28,585		20,450	

If we simply ignore the mobility by deleting mobile students, we get an estimate of the variance between schools that is larger than for all other models (Model 6 σ_u^2 = 2.500). A closer inspection of our data gives a possible explanation for this large increase in school level variance. Comparing the dataset of the delete approach with the dataset of the first school or last school approach, one can see a large reduction in number of students and number of dropouts, which has a greater impact in schools with a significant mobility. In the first school approach, e.g. school ID 66,628 has 13 dropouts on a total of 117 students; deleting the mobile students in this school, reduces the number of dropouts to 0 on a total of 58 remaining students. Deleting the mobile students has some consequences for the overall dataset: schools with a high (outgoing) student mobility, tend to have fewer dropouts, since dropout is related to changing schools. Therefore, differences in dropouts between schools become greater, resulting in an increase in school-level variance. Because we explicitly want to account for mobile students in our dropout study, we drop this approach from further analysis.

The last school approach and the cross-classified approach outperform the first school, the multiple membership approach and delete approach. Therefore, we will base our further comparisons of hierarchical and non-hierarchical models only on the last school and cross-classified approach. Concerning the research questions on dropout, we add student and school characteristics to the cross-classified baseline model, because this model yielded the best DIC-value.

Student and school characteristics

Starting from model 7 until model 9, we add student characteristics to the cross-classified model and evaluate the model fits. To facilitate the interpretation of the continuous variables, these variables were centred around their grand mean value. The effect of the predictor variables on the odds to dropout, is expressed as the percentage increase or decrease in the odds due to a one-unit increase, i.e. a one standard deviation increase in the predictor variable (Pampel, 2000). First, we add two variables referring to grade retention in the past school career, without controlling for other background characteristics. These two variables, RET_P ($p < .001$) and RET_S ($p < .001$) both have a significant effect on later dropout. This age of entry in secondary school seems to have a very strong impact on the chance for dropout: with an odds-ratio of 4.03, the odds of leaving secondary education without diploma for students who were older at the start, were about four times higher compared to students with the common age of entry. Also grade retention in secondary education (independent of grade retention in primary education) is a strong predictor of future dropout. The effect of grade retention on dropout is shown in the odds-ratio of 2.67, indicating that for students who repeated a grade, the odds of dropping out are 165% higher compared to students who never repeated a grade. Inclusion of the remaining student characteristics, GEN, COGN and SES, further lowers the DIC-value, indicating a better model fit (DIC = 3,967.253). Adding these covariates lowers the odds of age and grade retention slightly, but they remain significant predictors for future dropout. The effect of gender on dropout is revealed in several studies and replicated in this study: boys have a significant higher chance for dropout, compared to girls. With an odds-ratio of 0.60, girls have 40% lower odds to dropout, compared to boys. COGN has a significant effect on dropout ($p < .001$), whereas a one standard deviation increase in initial cognition (SD = 0.62075), lowers the odds for dropout with 38%. A one standard deviation increase in SES (SD = 2.17), lowers the odds for dropout with 47%. The addition of these student characteristics already explains a substantial part of the differences between schools and we can state that the chance for dropout is strongly defined by intake characteristics and grade retention in secondary education, even before incorporating school characteristics.

In an attempt to explain differences between schools in dropout rates, we also add school characteristics to our model. Four school characteristics are selected for our model. The best fit is obtained by adding the SES-S and RELAT-S variables. Including the COGN-S and GEN-S results in higher DIC-values and therefore, these variables are omitted from the final model. As can be found in Table 4, SES-S ($p < .001$) and RELAT-S ($p = .017$) have an effect on future dropout. Schools with higher mean socioeconomic composition, have far lower dropout rates. An increase of one standard deviation (SD = 0.46152) in the mean socioeconomic composition of the school, lowers the chance to dropout in that school with approximately 44.4%, even after controlling for student SES and other background characteristics. An increase of one standard deviation (SD = 0.15327) in RELAT-S, reduces the risk to dropout in that school with 21%. With the addition of these two school level variables, nearly all the between school variance can be explained.

Table 4: Cross-classified models with student and school characteristics

	Model 7		Model 8		Model 9		Model 10	
	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Constant	-8.529	*0.924*	-8.532	*0.623*	-9.292	*0.813*	-9.309	*0.640*
Grade 8	3.181***	*0.905*	3.127***	*0.625*	3.574***	*0.787*	3.110***	*0.631*
Grade 9	3.596***	*0.899*	3.637***	*0.619*	4.075***	*0.788*	3.600***	*0.626*
Grade 10	3.857***	*0.898*	3.950***	*0.617*	4.403***	*0.790*	3.912***	*0.624*
Grade 11	3.978***	*0.898*	4.114***	*0.618*	4.574***	*0.792*	4.078***	*0.624*
Grade 12	3.228***	*0.905*	3.349***	*0.628*	3.816***	*0.795*	3.317***	*0.632*
RET_S	0.977***	*0.099*	0.907***	*0.099*	0.916***	*0.103*	1.418***	*0.155*
RET_P	1.395***	*0.095*	0.978***	*0.103*	0.987***	*0.112*	1.766***	*0.148*
GEN			-0.536***	*0.111*	-0.485***	*0.114*	-0.440***	*0.110*
COGN			-0.720***	*0.082*	-0.692***	*0.087*	-0.669***	*0.080*

Continues on the next page

Table 4 (Continued)

	Model 7		Model 8		Model 9		Model 10	
	Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
SES			-0.297***	*0.034*	-0.276***	*0.037*	-0.670***	*0.068*
SES-S					-1.394***	*0.289*	-1.269***	*0.283*
RELAT-S					-1.553**	*0.629*	-1.419*	*0.604*
RET_P*SES							0.494***	*0.068*
RET_S*SES							0.325***	*0.071*
Between schools σ_u^2	1.245	*0.399*	0.369	*0.144*	0.166	*0.076*	0.143	*0.065*
DIC	4,118.007		3,967.253		3,929.775		3,881.30	
N	28,585		28,585		28,585		28,585	

Assumptions of discrete-time hazard models

Before drawing conclusions on the final model, we verify three important assumptions of the discrete-time hazard model: the proportionality assumption, the linearity assumption and the no unobserved heterogeneity assumption. Testing the proportionality assumption, one tests whether each predictor in the model has an identical effect in every time period under study. This is a very strong assumption and Singer and Willett (1991) argue that "violations of the proportionality assumption are the rule, not the exception" (p. 279). This assumption can be assessed by implementing interaction terms between the predictor and the time-variables (Grade). In our study, every interaction between the time variables and predictors was considered, but resulted in increasing DIC-values (not reported), indicating poorer model fits. Therefore, there is no reason to believe that the proportionality assumption has been violated. The linear additivity assumption, similar to the linearity assumption in linear regression, implies that a predictor's effect does not depend on the value of another predictor in the model. This assumption can easily be tested by looking for interactions between substantive predictors, and comparing the resulting model fits. In this study, only the SES-variable interacted significantly with both indicators of grade retention (RET_P and RET_S). Both interactions indicate a decreasing probability of dropping out for students with higher SES. However, this decrease is smaller when students are over-aged or retained in grade during secondary education. Both interactions are included in Model 10 and presented in Figure 2 and Figure 3.

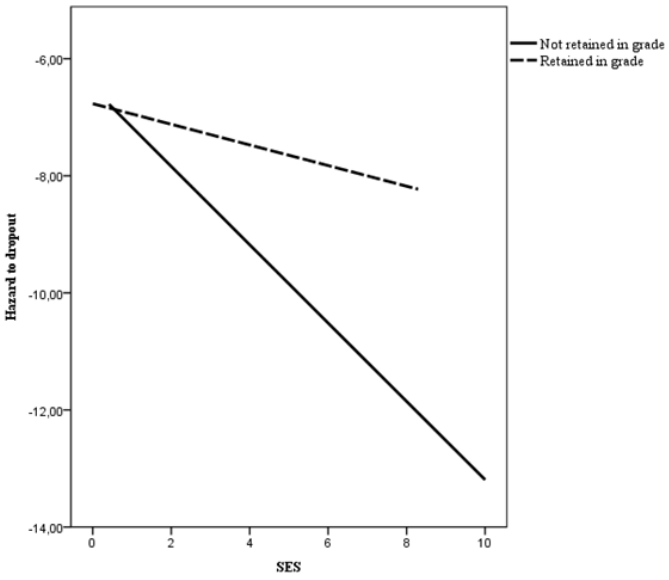


Figure 2: Interaction-effect socioeconomic status and retention in primary education

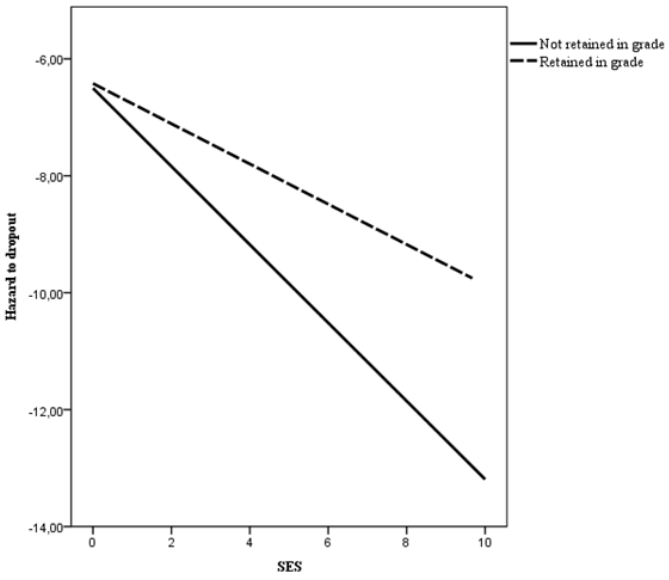


Figure 3: Interaction-effect socioeconomic status and retention in secondary education

The assumption of no unobserved heterogeneity, meaning that one or more important predictors have been omitted from the model, is the most difficult to assess. For checking this assumption, Singer and Willett (2003) recommend focusing on the hazard function, since unobserved heterogeneity has a consistent effect on the time variables and will lead to decreasing hazard functions. In the present study, the hazard functions increase over time (with a slight decrease in Grade 12), from which we can conclude that unobserved heterogeneity seems unproblematic.

Table 5: Impact of school characteristics in last school and cross-classification approach.

	Model 11 Last school		Model 12 Cross-classification	
	Est.	*SE*	Est.	*SE*
Constant	-9.410	*0.670*	-9.309	*0.640*
Grade 8	3.212***	*0.661*	3.110***	*0.631*
Grade 9	3.766***	*0.656*	3.600***	*0.626*
Grade 10	4.092***	*0.655*	3.912***	*0.624*
Grade 11	4.268***	*0.655*	4.078***	*0.624*
Grade 12	3.514***	*0.663*	3.317***	*0.632*
RET_S	1.416***	*0.155*	1.418***	*0.155*
RET_P	1.772***	*0.148*	1.766***	*0.148*
GEN	-0.445***	*0.110*	-0.440***	*0.110*
COGN	-0.689***	*0.080*	-0.669***	*0.080*
SES	-0.677***	*0.068*	-0.670***	*0.068*
SES-S	-1.101***	*0.275*	-1.269***	*0.283*
RELAT-S	-1.140	*0.582*	-1.419*	*0.604*
RET_P*SES	0.490***	*0.068*	0.494***	*0.068*
RET_S*SES	0.316***	*0.071*	0.325***	*0.071*
Between schools σ_u^2	0.125	*0.062*	0.143	*0.065*
DIC	3,904.416		3,881.30	
N	28,585		28,585	

Effect of cross-classification on the final model

Analogous to the comparison of the different baseline models, we compare estimates and variance components of the cross-classified approach with those of the last school approach (Table 5). Additionally, we focus on the effect of ignoring the cross-classification on the standard errors of the estimates.

Comparing the last school approach with the model with cross-classifications, differences between these two models on the standard errors arise, similar to those in Luo and Kwok (2009). Compared to the cross-classified approach (Model 12), the standard errors associated with the time variables of the last school approach (Model 11) are substantially higher. In contrast, the standard errors associated with the school characteristics in the last school approach are lower, while the standard errors associated with the student characteristics remain unaffected. Following Luo and Kwok (2009), this was somewhat expected because ignoring a cross-classification results in an overestimation of the standard errors at the (*k*-1)th level and an underestimation of the standard errors of the estimates of the variables related with the ignored crossed factor (at the *k*-th level). The standard errors associated with the predictor variables of the remaining crossed factor at the *k*-th level, are found to remain stable. Applied to our last school approach, we ignore the fact that grades at Level-1 are crossed with students and schools at Level-2. Ignoring this crossing at Level-2, resulted in higher standard errors of the regression coefficients of the predictor variables on Level-1 (Grade) and lower standard errors of the variables related with the ignored crossed factor (Schools). The standard errors associated with the student characteristics remain unaffected. The underestimation of standard errors can become problematic in terms of significance of the associated predictor, and can therefore lead to different conclusions. The variance components of our final model also differ depending on the classification. Ignoring the cross-classification affects the school level variance in a similar way as for the standard errors: the between school variance component in the last school approach seems to be an underestimation of the actual between school variance.

Comparing the effects of school characteristics on the last school model and the cross-classified model emphasizes the importance of incorporating student mobility in dropout research. In the last school approach, a school with a socioeconomic composition (SES-S) one standard deviation (SD = 0.46152) above the mean, reduces the probability to dropout in that school by 39%, which is a smaller impact compared to the cross-classified approach. However, the effect of perceived relationships at school (RELAT-S) indicates the importance of including student mobility. In the last school approach, this RELAT-S becomes non-significant, whereas in the cross-classified approach, RELAT-S is a significant predictor of future dropout.

Besides the bias in standard errors and the non-significance of school level variables, also the lower DIC model fit indicates the superiority of the cross-classified model.

DISCUSSION

Research on dropout in secondary education has been the topic of many studies and remains a contemporary topic. The interest for this dropout phenomenon is obvious: not only do dropouts have lower job opportunities (Solga, 2002), when they do find a job, these jobs are characterized by a lower social status and lower earnings (Verhofstadt, De Witte, & Omey, 2007). These strong implications justify the permanent attention of researchers for dropout. However, past research has some methodological shortcomings. In this study, we attempted to address some of these shortcomings by adopting a multilevel discrete time hazard model. Although the hazard approach is the most appropriate method in studying dropout and was implemented in some dropout-research (e.g. Bowers 2010; Gesthuizen, De Graaf, & Kraaykamp, 2005; Roderick, 1994), past research which adopted this method ignored the hierarchical structure of educational data. When a multilevel approach was implemented, researchers did not correct for student mobility despite the strong association with student dropout.

This discussion will focus on the two main topics of this study. At first, we discuss the results of the different approaches in modelling dropout and student mobility. The second part gives an overview of the results regarding the student and school characteristics predicting dropout in secondary education.

Traditionally, past research on dropout, which implemented a hazard analysis, ignored the fact that students are nested within schools. Replicating this approach, and comparing this approach with multilevel approaches, this study shows an overestimation of the time variables and a higher DIC-value indicating a poorer fit. In a next step, we classify students on different school ID's, from which different results arise. In line with Grady and Beretvas (2010), models which are based on the school ID of the first school attended, yield the highest model fits and lowest between-school variance, indicating that this approach fails to explain some part of the variation in the data and should be avoided in future research. Their study, however, was limited in the fact that they did not focus on the last school approach. For dropout research, taking the ID of the school at which the student dropped out, seems more rational compared to the first school approach. Results of

the last school approach indicate that, for this dataset, classifying students on the last school attended produces better fits and results in substantially higher between-school variance. Extending these approaches and taking into account the student mobility by means of a cross-classification with measurement occasions nested within a crossing of students and schools at Level-2, results in a better fit compared to the last school approach, which underlines the importance of incorporating the mobility. If we accounted for the student mobility by means of a multiple membership model, this resulted in a higher DIC-value compared to the last school approach, which was somewhat unexpected based on previous research. The poorer model fit stresses the importance of the choice of correct weighting schemes. In this study, we use the proportional time spent in every school, which seems, compared to the model fits of the cross-classified and last school approach, an oversimplification of the school effects on dropout. It seems to be the case that not every school attended has the same impact, and comparing the model fits of the last school, cross-classification and multiple membership approach indicates that more recent schools have a greater impact, as suggested by Fielding and Goldstein (2006). As mentioned, not only the time spent in a particular school is of importance, but also the timing and order the student was in different schools.
By way of illustration (and because of common practice), we also estimate a model where we delete school changers. This deletion results in a reduction of our dataset of about 30% of the subjects, which makes a comparison with the previous approaches impossible.

After the addition of student and school characteristics to the model, the impact of incorporating student mobility becomes more pronounced. We added these student and school characteristics to the cross-classified model and compared this model with the last school approach, since these two approaches yield the best model fits. In line with Luo and Kwok (2009), the standard errors associated with the time variables in the last school approach seem to be overestimated, in contrast to an underestimation of the standard errors of the parameters associated with the ignored crossed factor. More important is the impact on the fixed effects parameters. Where parameters associated with student characteristics remain relatively stable, parameters associated with school characteristics and the time

variables change to a large(r) extent. Results indicate a similar underestimation of school characteristics for the last school approach, which led to insignificant school characteristics compared with the cross-classified approach. In general, these school level variables seem to have a stronger impact on the risk to dropout than generally assumed. This difference emphasizes the importance of adopting a cross-classified approach in studying dropout.

The conclusions regarding dropout in secondary education are based on the cross-classified approach. The highest chance to dropout is in Grade 11, with a decrease in Grade 12. The highest chance in Grade 11 is not surprising since a lot of students who dropout were retained in grade and compulsory education ends at the age of 18. Student characteristics predicting dropout are very similar to previous research (Alexander et al., 2001; Lee and Burkham, 2003; Marks, 2007). Our analyses confirm the negative impact of grade retention on dropout, as found in previous studies. Results indicate that grade retention is a strong indicator of future dropout, irrespective of the fact the student was retained in grade in primary or secondary education. This impact of grade retention remains, even after controlling for other relevant student background characteristics. In line with Alexander et al. (2001), we conclude that early achievement can also predict future dropout, even after correcting for grade retention in primary and secondary education. As for the effect of gender, our results are in line with e.g. Marks (2007): boys tend to have higher odds of leaving school without a diploma, compared to girls. The effect of the final student characteristic, socioeconomic status, was in line with previous research and highlights the impact of SES on dropout. Children stemming from a lower socioeconomic class, tend to have higher chances to dropout irrespective of their achievement. The final model in this study also presents an interaction between SES and the two variables related to grade retention, where the effect of SES on future dropout is different depending on the fact the student was retained in grade or not. Students retained in grade during primary education, tend to have high chances to dropout, regardless their SES-status. On the other hand, students who started secondary education on time, but are from a low SES-family, have a very high chance to dropout, compared to students from high SES-families. If a student was retained in grade during secondary education, the SES-status of the student will

determine if a student drops out. Retained students with a high SES tend to have lower chances to dropout compared to retained students with a low SES.

Regarding the effects of schools on dropout, both the socioeconomic composition and the relationships at school seem to predict the chances to dropout. In line with Rumberger and Thomas (2000), we conclude that high-SES schools have lower dropout rates compared to schools with an average SES-composition. We can draw similar conclusions regarding the effect of school climate on the chance to dropout: schools where the relationships between students and teachers are better, tend to have lower dropout rates.

CONCLUSION

With this study, we attempt to come towards some methodological shortcomings in previous dropout research. We implement a multilevel discrete-time hazard model, in which we correct for school changes by adopting a multiple membership approach and a cross-classified approach. These approaches are demonstrated on a real dataset. We conclude that ignoring the hierarchical nature of educational data, and ignoring student mobility can lead to different conclusions on the predictors of dropout. With this discrete-time approach, we are also able to draw conclusions on the moment of dropout and the implementation of time-varying variables becomes straightforward. Overall, this study contributes to the modelling of dropout, accounting for student mobility.

Chapter 3

Different pathways towards dropout: the role of engagement in early school leaving

Published as: Lamote, C., Speybroeck, S., Van Den Noortgate, W., & Van Damme, J. (2013). Different pathways towards dropout: the role of engagement in early school leaving. *Oxford Review of Education, 39(6)*.

ABSTRACT

In this study, we examine the development of student engagement in relation to dropout. We focus on different growth trajectories of engagement between groups of students and on whether these trajectories lead to differences in the survival of the student. The development of behavioural and emotional engagement of 4063 graduates and 541 (11.7%) dropouts is examined from Year 7 through Year 12 and this development is linked to the probability of dropping out in each grade by means of a discrete-time survival mixture model. For emotional engagement, results point to a model with two different subgroups: one group starting at a high level of engagement and following a (relatively) stable pattern and the other group starting at a lower level of engagement and following a decreasing trend. For behavioural engagement, the results indicate that a 3-class model showed the best fit: a high and (relatively) stable group, a high and decreasing group and a low and stable group. In terms of dropout, the unstable and low groups demonstrate a significantly higher probability of dropping out, as evidenced in the steep, declining survival curves. Different background variables are included to gain more insight into engagement and dropout, and to predict membership in the low and decreasing class.

INTRODUCTION

In most countries, students who complete secondary school and receive their diploma/certification encounter fewer problems in later life compared to students who leave school before receiving their certification (dropouts). For instance, qualified students have better job opportunities (Solga, 2002) and are less likely to be employed in poorly paid jobs (EACEA Eurydice, 2012). Therefore, preventing students from leaving school before they obtain their diploma/certification remains important.

Previous research on predictors of dropping out primarily focuses on the 'usual suspects', that is: a set of recurring variables that includes gender, grade retention, socio-economic status, ethnicity and achievement. However, school dropout can also be linked to a decline in school engagement. If a student feels more engaged in school, he is more likely to perform better and is less at risk of dropping out (Alexander et al., 2001; Fredricks et al., 2004). When engagement was previously integrated in dropout research, only one or two measurement moments were used and engagement was measured as a one-dimensional construct; additionally, only a few studies took the multidimensional and longitudinal nature of engagement into account. In this study, we focus on the predictive power of engagement on dropping out. When we focus on this engagement, we consider the longitudinal and multidimensional nature of engagement, and we relate this to the probability of dropping out in every grade.

We examine the situation in Flanders (Belgium). Currently, Flanders has a dropout rate of approximately 14%, which significantly exceeds the Europe 2020 benchmark of 10% (European Commission, 2010). Therefore, as in several other European member states, reducing early school leaving in Flanders is a key challenge they face in the years to come.

THEORETICAL FRAMEWORK

School engagement

School engagement has been investigated in numerous studies (for an overview, see e.g. Christenson, Reschly and Wylie (2012)) because of its strong positive association with both learning outcomes and social outcomes (Klem & Connell, 2004).The interest in school engagement is hardly surprising given that its malleable nature easily lends itself as a target for interventions (Fredricks et al., 2004). The concept of school engagement, however, has been defined in various ways. In his well-known participation-identification model, Finn (1989) refers to engagement as a two-dimensional construct with a behavioural component (participation) and an affective component (identification). Both dimensions are linked to each other and to school outcomes. The identification with school is influenced by successful school outcomes; students with lower school outcomes (e.g. shown in grade retention) tend to show less identification with school. This lower level of identification influences the participation in curricular and extracurricular activities and can potentially result in increased levels of truancy. This reduced participation, in turn, results in even lower levels of academic achievement. The behavioural and affective component were also a part of the definition of school engagement of Kortering and Christenson (2009). They defined school engagement as "…a concept that requires psychological connections within the academic environment […] in addition to active student behaviour" (p. 7) wherein 'psychological connection' and 'active student behaviour' can be interpreted as equivalent to the identification and participation components of Finn's model.

In contrast with the former authors, Fredricks et al. (2004) distinguished between three dimensions of engagement: behavioural, emotional and cognitive. Behavioural engagement refers to involvement in academic and social activities; emotional engagement refers to the relations with teachers, classmates and/or school; and cognitive engagement refers to the willingness to put effort in complex problems and tasks. This three-dimensional structure of school engagement is

widely accepted and has been validated in research on student engagement (Archambault et al., 2009; Wang, Willett, & Eccles, 2011).

Self-concept is often posited as an important antecedent of school engagement. In their 'Process Model of Perceived Control and School Performance', Skinner, Wellborn and Connell (1990) find that a student's perceived control has a positive effect on engagement. More specifically, a positive self-concept of ability results in higher levels of task engagement (Eccles, 1983). This relation is also a part of the conceptual model of high school performance of Rumberger and Rotermund (2012), where self-perception (as an overall term for self-concept and self-esteem) is considered as a precursor of engagement and engagement is viewed as precursor of student achievement, which can in turn influence subsequent attitudes, including self-perception.

School engagement and dropout

School engagement has been the central theme in most dropout theories (e.g. Finn, 1989; Rumberger & Larson, 1998). According to these theories, the decision of students to drop out is a consequence of a gradual disengagement from school, either for social or academic reasons (Rumberger, 1987). Rumberger emphasises that to develop a more comprehensive model for understanding dropout, one should focus on processes underlying the decisions to drop out rather than only focussing on the direct factors related to this dropping out (e.g. socio-economic status). However, only a few studies have considered the longitudinal nature of the engagement process when examining determinants of the student's decision to drop out.

Apart from this developmental aspect, one should consider the multidimensionality of engagement. Fredricks et al. (2004) notes that most of the research on dropout has focussed on the behavioural component, and thus, there is less empirical evidence on the relation between emotional engagement and dropout. Additionally, the cognitive engagement is the most difficult one to assess and integrate in research.

More recently, some studies have attempted to meet some of these shortcomings by explicitly linking dropout with the various engagement dimensions. Pannozzo, Finn and Boyd-Zaharias (2004, in Finn & Zimmer, 2012) include measurements of behavioural and affective engagement in grade 4 and grade 8 into one model and find effects of engagement in both grades on dropout. Accordingly, declines in both behavioural and affective engagement contributed to the decision to withdraw from school. Although Pannozzo includes measures of engagement at two different points in time, the longitudinal character of disengagement could be considered problematic given that only determining the difference between two time points is insufficient to adequately describe change (Singer & Willett, 2003).

On the basis of a 3-wave measurement design, Archambault et al. (2009) constructs individual school engagement trajectories, and accordingly, the authors identify six different trajectories. These trajectories differ from the beginning level of engagement at age 12, in the final level at age 16, and in the stability of engagement over time. These different trajectories reveal different probabilities for dropout as students with unstable engagement pathways demonstrate significantly higher odds to dropout, than do students with stable pathways. A similar conclusion was drawn by Janosz et al. (2008). Although both studies consider the longitudinal nature of engagement, they do not relate this with a survival analysis of dropout and thereby ignore the timing of dropout.

The present study aims to test the impact of the development of school engagement on the probability of dropping out. To adequately model the longitudinal effect of school engagement on dropout, we use a discrete-time survival mixture model (DTSMA; Muthén & Masyn, 2005). Three research questions guide our study: (1) Can we define subgroups of students based on different trajectories of school engagement during secondary education? (2) Do different trajectories reflect different probabilities related to dropping out , and timing of dropping out ? (3) Which factors predict class membership?

METHOD

Participants

For this study, the data were drawn from the Flemish 'LOSO' project (Van Damme et al., 2002). This longitudinal research project was initiated in 1990, and it followed a cohort of 6,411 students throughout their secondary education, with a follow-up on the labour market or in higher education. Of the 6,411 students, we selected a subsample of 4,604 students from 52 schools. The number of students per school varied from 5 to 347. More specifically, we selected all students who started in the first grade of mainstream secondary education and had known 'educational positions' (Grade and school ID's) for every school year through secondary education. Students who were lost to follow-up were excluded. Additionally, because we account for the multilevel nature of the data, with students nested in schools (see Analysis-section), we excluded students who moved during the study to a school that was not included in the LOSO-study. From these 4,604 students, 541 (11.8%) were identified as dropouts. This rate is somewhat smaller than the overall Flemish dropout rate of 14%, because students from special education were not considered in our total group (Van Landeghem & Van Damme, 2011b).

Measures

Dropout

A student was considered a dropout, the moment he leaves fulltime mainstream secondary education before having obtained the degree of higher secondary education. This means that students leaving full-time education and entering into part-time education (or special education) were also considered as having dropped out[vi]. Although they are able to obtain specific types of qualifications, previous research in Flanders has shown that these qualifications do not guarantee the same outcomes in the labour market or in higher education (Creten et al., 2004)[vii], compared to a diploma of secondary education. Generally, students participating in part-time education obtain a qualification that is ranked lower than the secondary education diploma on the Flemish Qualification Structure (derived from the

European Qualification Framework, EQF), or they do not obtain a qualification at all.

Engagement

A well-being questionnaire was administered to the students at the end of Year 7, 8, 10 and 12. The 8 original scales were adopted from the Schoolvragenlijst Voortgezet Onderwijs (School Questionnaire of Secondary Education) of Smits and Vorst (1982) and were confirmed in the LOSO-study with factor analyses on data collected four moments. The scales all yielded adequate internal consistencies (Cronbach's $\alpha > .80$) (see: Van Damme et al., 2002). In accordance with Fredricks et al.'s (2004) definition of school engagement , we selected for this study two scales of this questionnaire: 'relationship with teachers' as an indicator of emotional engagement and 'attitude toward homework' as an indicator of behavioural engagement (for specific items: see Appendix 1) (Van de gaer et al., 2009). Unfortunately, we have no indicators for cognitive engagement at our disposal.

Background variables

Apart from the dropout-status and engagement scores, we included 7 student background variables: gender, socio-economic status (SES), initial cognitive ability, age at the start of secondary education, grade retention in secondary education, start in a remedial class and ethnicity of the student.

Gender is expressed as a dummy variable, with male students as the reference category (male = '0', female = '1'). The continuous variable referring to the SES of the student is based on parental occupational status, parental education, cultural capital and monthly income of the family. Combining these variables into a factor analysis yielded a one factor solution, which represents a score for SES. The variable 'initial cognitive ability of the student' stems from a factor analysis of 3 measures: math achievement, language achievement and an intelligence score, all of which were measured at the start of Year 7 (Van Damme et al., 2002). The age at the start of secondary education was used as an indicator for grade retention in primary education ('0' = normal age of 12 or younger, '1' = older). Grade retention in secondary school was implemented as a dummy variable with a value of '1' for

students who were retained during secondary education. The next background variable, start in a remedial class, refers to a specific aspect of the Flemish educational system. Students who were not able to meet the final attainment level of primary education and /or did not obtain a certificate of primary education are assigned to a remedial class. After a year in the remedial class, students can switch over to the first year of the general secondary education or to the second pre-vocational year. The last background variable is the ethnic origin of the student, with migrant students coded as '1' (Lacante, Almaci, Van Esbroeck, Lens, & De Metsenaere, 2007).

Analysis

For this study, we use Mplus 6.12 (Muthén & Muthén, 2010) to conduct a discrete-time survival mixture analysis (Muthén & Masyn, 2005). The combination of discrete-time survival analysis with a growth mixture model captures unobserved class membership by modelling the association between different trajectories of engagement and the risk of dropping out. In other words, we can examine whether students differ on their growth trajectory of engagement and whether these different trajectories lead to different probabilities that the student will dropout.

The growth mixture aspect of the model is similar to a traditional latent growth model, but allows for differences in growth parameters across different unobserved subgroups. In traditional latent growth models one assumes that, possibly after correction for known background variables, one growth pattern holds for all cases, whereas different growth patterns for different unknown subgroups may exist. To capture these subgroups, while a cluster analysis can be conducted, the growth mixture approach has several advantages over the cluster analysis. One major advantage is that mixture modelling is a 'model-based cluster analysis', in which the obtained clusters follow from looking for similar patterns of growth. Estimating this mixture model results for each person in posterior probabilities of group memberships and based on the highest posterior probability, an individual is assigned to a certain cluster. A second advantage is that, with a mixture model, we are able to relate covariates to the probability of group membership. A more extensive overview of advantages of this mixture approach, compared to a cluster

analysis can be found in McLachlan and Chang (2004) and Nagin (1999). In this study, trajectories for the two selected indicators of engagement were estimated in two separate analyses.

The survival part of the model describes the probability of dropout for every time period. Such a survival analysis is most appropriate in analysing the occurrence of an event in a certain time period, starting at a well-defined starting point (Year 7) and ending at the occurrence of the event of interest (dropout) or completion. A discrete time survival analysis (DTSA; Singer & Willett, 2003) has proven its value in analysing dropout in secondary education (Bowers, 2010; Gesthuizen et al., 2005; Lamote, Van Damme, Van Den Noortgate, Speybroeck, Boonen, & de Bilde, 2013; Roderick, 1994). The discrete-time aspect follows from the reality of education, in which the majority of students drop out at the end of a school-year (or the beginning of a subsequent school-year), while only a few students dropped out in the middle of a school-year. Therefore, it seems reasonable to treat the moment of dropout as a discrete variable.

Generally, the DTSA is conducted in a logistic regression framework. However, this discrete-time analysis can also be modelled in a more complex latent class regression framework and linked to a mixture component, in which survival is predicted by growth trajectory classes, thus resulting in a discrete-time survival mixture analysis (DTSMA). When linking this mixture model with a survival model, we assume an impact of different engagement trajectories on the grade level of student dropout. This model is depicted in Figure 4: Discrete-time survival mixture model, with time-invariant covariates (McCullough, Friedman, Enders, & Martin, 2009; Muthén & Masyn, 2005).

In Figure 4, the rectangle with the x represents all of the covariates included in the final DTSMA model, the ellipse c the latent trajectory groups, the ellipse η_y the latent growth model of engagement and the ellipse η_u the latent discrete-time survival model. The arrow from c to the survival part η_u indicates that the survival parameters can differ between classes, and the arrow from c to the growth part η_y indicates that the means of the growth factors can vary across classes. The arrow from x to the survival part η_u represents the effect of the

covariates on the survival of the student, and the arrow from *x* to *c* represents the multinomial logistic regression of *c* on *x*, the latent class regression, which is a description of the relation between the covariates and the class membership.

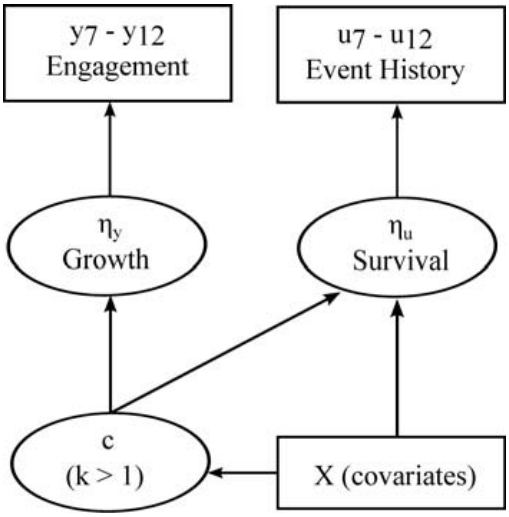


Figure 4: Discrete-time survival mixture model

The data used in this study are multilevel in nature, with students clustered within schools. The clustering was based on the student's last school attended. To account for this clustering, standard errors of parameter estimates were estimated using a 'sandwich estimator' that assumed independent observations only across schools, and a robust maximum likelihood estimator was used to account for non-normality of the data. (Muthén et al., 2002). The evaluation of the model fit was based on the Lo-Mendell-Rubin Adjusted Likelihood Ratio Test (aLMR) and on the Bayesian Information Criterion (BIC) (Nylund, Asparouhov, & Muthén, 2007). For correct class enumeration, we did not estimate the model without the covariates as previous research has clearly shown that a mixture model specified without covariates, is very likely to be miss-specified (Li & Hser, 2011; Muthén, 2004). Therefore, a model with only intercept and slope, that is, without covariates, is not estimated/reported. The reported tables should be interpreted as the results of one model estimation (but with separate analyses for the different engagement dimensions).

RESULTS

Dropout

Of the 4,604 students starting their first year of secondary education, 541 students dropped out and did not obtain a higher secondary education diploma (Table 6). The hazard h to dropout, i.e. the conditional probability that a student will experience the event of dropping out in a certain grade (given that they did not experience the event in an earlier time period) increased beginning in Year 7 and peaked at Year 11, followed by a small decrease in Year 12. At the end of secondary education, more than 88% of the students did not drop out and obtained a diploma of higher secondary education.

Table 6: Life Table

Year	Enrolled at beginning of year	Not in subsequent year	Censored	Hazard *h*	Survival *s*
0	0	0		0.0000	1.0000
7	4,604	3		0.0007	0.9993
8	4,601	54		0.0117	0.9876
9	4,547	95		0.0209	0.9670
10	4,452	147		0.0330	0.9351
11	4,305	151		0.0351	0.9023
12	4,154	91	4,063	0.0219	0.8825

Discrete-time survival mixture analysis

We estimated a separate DTSMA for behavioural and emotional engagement and thus report the results separately. For every dimension of engagement, we first discuss the number of classes and the growth parameters of the different classes. (i.e. in Figure 4, the arrow connecting *c* with the growth part η_y). Second, we examine whether class membership has an effect on the hazard to drop out (i.e. in Figure 4 the arrow from *c* to η_u). Third, we examine the effect of the covariates on the survival of the student (i.e. in Figure 4 the arrow connecting *x* to the survival part η_u). Fourth, we use these covariates to predict class membership (i.e. in Figure 4 the arrow from *x* to *c*, the latent class regression). Finally, and perhaps one of the main reasons to conduct a DTSMA, we describe the effect of class membership on survival.

Behavioural engagement

For behavioural engagement, a 3-class linear model yielded the best aLMR and BIC scores, indicating the best model fit for this dataset. This 3-class model (Table 7) consisted of a trajectory with students starting at a high level of engagement (M = 4.398) and following only a small decline during secondary education (= 'High group'). A second group of students started at almost the same level of engagement (M = 4.346) compared to the 'High group', and the difference in intercepts between these two trajectories was not significant ($\chi^2(1) = 1.540$, $p = .21$). However, these students followed a significantly steeper decline in engagement than the 'High group' (-0.312 versus -0.092, $\chi^2(1) = 153.67$, $p < .001$). We label this steep declining group as 'High and Decreasing group'. The last group of students started at a significantly lower level of engagement compared to the 'High group' (χ^2 (1) = 866.574, $p < .001$) and remained at this level during their time in secondary education (='Low group'). The differences between these three classes are illustrated in Figure 5, plot A.

Table 7: Parameter estimates for growth parameters of behavioural engagement

Parameter estimates	High & Decreasing		Low		High	
	Est.	*SE*	Est.	*SE*	Est.	*SE*
Means						
Intercept	4.346***	*0.048*	3.263***	*0.042*	4.398***	*0.018*
Linear Slope	-0.312***	*0.022*	-0.016	*0.010*	-0.092***	*0.006*
Class proportions	*0.13*		*0.21*		*0.66*	

Entropy 0.635
Note. * *p* < .05, ** *p* < .01, *** *p* < .001

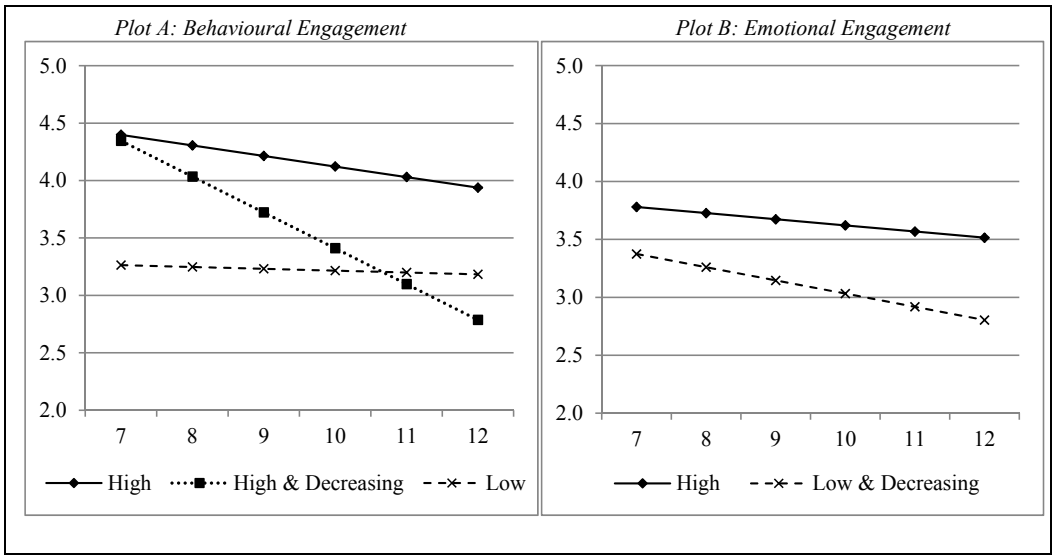


Figure 5: average change trajectories of (a) behavioural engagement and (b) emotional engagement

More important in this DTSMA are the survival parameters in Table 8. The intercept of this survival part is an estimate of the influence of the latent class on the hazard to dropout compared to a reference class (the 'High group' in this study). The significance of this estimate indicates an effect of class membership on hazard, with students in the 'Low' and 'High and decreasing' classes having a greater hazard to dropout. The predictors of dropout indicate that students starting with a low initial cognitive ability, starting in a remedial class, having an ethnic background, having a low socio-economic background, being older at the start of secondary education and/or being retained in grade in secondary education are significantly more at risk for dropping out. The effect of gender is not significant.

Table 8: Behavioural engagement: effects of covariates on survival

	High & Decreasing		Low		High	
	Est.	*SE*	Est.	*SE*	Est.	*SE*
Intercept	2.417***	*0.364*	2.315***	*0.361*	0.000	*fixed*
Initial Cogn. Level	-0.799***	*0.139*				
Start 1B	0.372***	*0.171*				
SES	-0.290***	*0.041*				
Gender	-0.035	*0.125*				
Grade retention SE	0.617***	*0.126*				
Age at start SE	0.720***	*0.157*				
Ethnicity	-0.376***	*0.191*				

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

Of particular interest are the regression coefficient estimates for predicting latent class membership (Table 9). We used the 'High group' as the reference category, and we report the probability of belonging to the 'High and decreasing group' and the 'Low group' compared to the 'High group' as odds ratios (with 95% confidence intervals)[viii]. Relative to the students in the reference category, students with a lower SES were more likely to belong to the 'Low group' (OR = .94, 95% CI .89 – .99). We found a similar relation between SES and the 'High and Decreasing' group, where students with a low SES have a greater likelihood to be member of this 'High and Decreasing' group (OR = .92, 95% CI .86 – .98). Male students were also more likely to be in the 'Low group' (OR = .19, 95% CI .15 – .24) and the 'High and Decreasing' group (OR = .36, 95% CI .27 – .48). The odds of belonging to the

'Low group' versus the 'High group' are 2.73 times (95% CI 1.52 – 4.91) higher for students starting in the remedial class compared to students starting in regular secondary education. The students who begin in the remedial class are also 2.21 (95% CI 1.31 – 3.72) times more likely to be in the 'High and Decreasing' group compared to being in the 'High group'. Students retained in grade in secondary education were more likely to be members of the 'High and Decreasing group' (OR = 4.57, 95% CI 3.42 – 6.11) or the 'Low group' (OR = 3.22, 95% CI 2.50 – 4.15) compared to the reference group.

Table 9: Behavioural engagement: Covariates predicting class membership.

	High & Decreasing		Low	
	Est.	*SE*	Est.	*SE*
Intercept	-0.619	*0.946*	0.473	*0.724*
Initial Cogn. Level	-0.081	*0.170*	-0.204	*0.131*
Start 1B	1.004***	*0.357*	0.792***	*0.317*
SES	-0.085*	*0.038*	-0.064*	*0.032*
Gender	-1.023***	*0.178*	-1.676***	*0.150*
Grade retention SE	1.520***	*0.176*	1.169***	*0.154*
Age at start SE	-0.025	*0.341*	0.283	*0.185*
Ethnicity	0.042	*0.265*	0.559***	*0.183*

Note. * $p < .05$, ** $p < .01$, *** $p < .001$

In terms of survival of these two groups, the estimated survival probability of the 'High group' follows a more stable pattern, compared to the decline in survival probability of the 'High and Decreasing group' and 'Low group' (Figure 6, plot A).

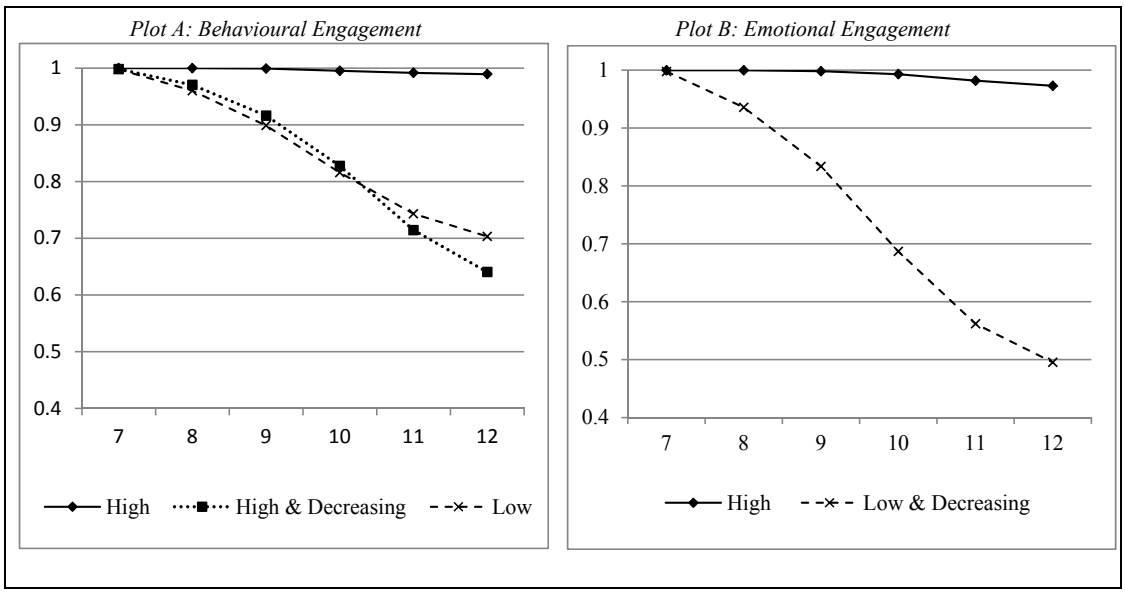


Figure 6: estimated survival probability of (a) behavioural engagement and (b) emotional engagement by grade.

Emotional engagement

With respect to emotional engagement, the aLMR and BIC indicated that a 2-class linear model was the best model for this dataset. These two classes had significantly different ($\chi^2(1) = 61.312$, $p < .001$) starting levels of emotional engagement in the first year of secondary education. Both groups followed a decreasing pattern, but this decline was significantly steeper for the group starting at the lowest level of emotional engagement (-0.053 versus -0.114; $\chi^2(1) = 4.612$, $p = .03$). These trajectories are depicted in Figure 5, plot B. We label the group starting at a higher level and following a more stable (less decreasing) pattern, the 'High group'. The group starting at a lower level and following the steeper decline is labelled as the 'Low and Decreasing group' (Table 10).

Table 10: Parameter estimates for growth parameters of emotional engagement

Parameter estimates	Low & Decreasing		High	
	Est.	*SE*	Est.	*SE*
Means				
Intercept	3.374***	*0.046*	3.780***	*0.028*
Linear Slope	-0.114***	*0.025*	-0.053***	*0.006*
Class proportions	*0.19*		*0.81*	

Entropy 0.605
Note. * *p* < .05, ** *p* < .01, *** *p* < .001*

Analogous to the latent class growth intercept of the behavioural engagement, the significance of the intercept of the latent class growth in emotional engagement in Table 11 indicates that students in the 'Low and Decreasing group' have a higher probability of dropping out[ix]. In addition to this class membership, dropout was predicted by being a low-performer at the start of secondary education, having a lower SES-score, being older at the start of secondary education and having repeated a grade during secondary education.

Table 11: Emotional engagement: effects of covariates on survival

	Low & Decreasing		High	
	Est.	*SE*	Est.	*SE*
Intercept	2.874***	*0.221*	0.000	*fixed*
Initial Cogn. Level	-1.017***	*0.165*		
Start 1B	0.314	*0.307*		
SES	-0.312***	*0.050*		
Gender	0.383	*0.217*		
Grade retention SE	0.428**	*0.190*		
Age at start SE	1.009***	*0.220*		
Ethnicity	-0.068	*0.201*		

Note. * *p* < .05, ** *p* < .01, *** *p* < .001

Class membership was evaluated by comparisons with the 'High group'. As evidenced in Table 12, the probability of being a member of the 'Low and Decreasing group' increases significantly when a student is male (OR = .20, 95% CI .13 – .31) or has a low socio-economic background (OR = .89, 95% CI .83 – .96). A student retained in grade during his secondary education is also more likely to be in the 'Low and Decreasing' trajectory compared to the 'High' trajectory (OR = 3.98, 95% CI 2.64 – 6.01).

Table 12: Emotional engagement: Covariates predicting class membership.

	Low & Decreasing	
	Est.	*SE*
Intercept	0.259	*1.106*
Initial Cogn. Level	-0.081	*0.170*
Start 1B	1.004	*0.357*
SES	-0.085**	*0.038*
Gender	-1.023***	*0.178*
Grade retention SE	1.520***	*0.176*
Age at start SE	-0.025	*0.341*
Ethnicity	0.042	*0.265*

Note. * *p* < .05, ** *p* < .01, *** *p* < .001

Regarding the differences in survival between the 'Low and Decreasing' and 'High group', the survival curve for the 'Low and Decreasing group' demonstrated a steeper decline, indicating a lower survival probability for students who start their secondary education at a lower level of emotional engagement and follow this declining pattern of emotional engagement (Figure 6, plot B).

DISCUSSION

In this study, we explored the dropout phenomenon in Flanders and the relation between dropout and the development of school engagement. We extended the study of Archambault et al. (2009) and Janosz et al. (2008) by combining a mixture model with a survival analysis. With this survival analysis, we were able to draw conclusions on the timing of dropout and we also attempted to address some critiques on previous research regarding dropout (e.g. Fredricks et al., 2004), by accounting for the longitudinal nature of dropout, for the possibility of multiple unknown groups, and for the relation with different kinds of engagement

Our main interests were two-fold. First, we wanted to determine whether the development of both engagement constructs is the same for every student or whether different subgroups of development of engagement existed. Second, we were interested in the relation between these subgroups and the timing of dropout. Do students in one subgroup drop out earlier compared to students in other subgroups? Additionally, we were interested in the background characteristics of the students in the different subgroups. In search for an answer to these questions, we combined a mixture analysis with a survival analysis, which allowed us to draw conclusions regarding the development of engagement and in the grade during which this dropout occurs.

Development of behavioural and emotional engagement

Following Fredricks et al. (2004), we acknowledged the multidimensional and longitudinal character of student engagement in considering the development of behavioural and emotional engagement. For both constructs, we found a group of students starting at a higher level and following a relatively stable pattern (with only a slight decrease) as they progressed through secondary education. Consistent with previous research on student engagement and dropout (Archambault et al., 2009; Janosz et al., 2008), this ‘normative’ group consists of more than half of the students for both engagement constructs.

More specifically, over 60% of the students feel behaviourally engaged, i.e. they demonstrate a positive attitude towards homework, and (more or less) stay

at this level. A minority of the students reported a high behavioural engagement in Year 7, but became more disengaged as they advanced in their secondary education. Considering the specific indicator of behavioural engagement used in this study, that is, attitudes towards homework, this decline can be expected as students mature (Xu, 2004; Van de gaer et al., 2009).

In contrast with Janosz et al. (2008), we found that not only the unstable pathway of engagement showed a higher hazard to dropout. In our study, about 20% of the students entered secondary education with a significantly lower level of behavioural engagement. These students stay at this level and have a significantly higher probability of leaving school without a diploma. This result corroborates some of the findings of Blondal and Adalbjarnardottir (2012) where the group of 'expected dropouts' was mainly characterised by low behavioural engagement (negative school behaviour). Similar to our results, Blondal and Adalbjarnardottir did not find any (positive or negative) changes in engagement for this group of disengaged students over time.

When examining the trajectories of emotional engagement, both trajectories showed a decline, although this decline is more pronounced for the group of students who started at a lower level. Previous research that focussed on the developmental aspect of emotional engagement (more specifically on student-teacher relationships) reported an overall declining quality of these student-teacher relationships (Van de gaer et al., 2009), even within one year (Opdenakker, Maulana, & den Brok, 2012). In our study, the declining trend continued through secondary education, with only a difference in the degree of decline between the subgroups. A possible explanation can be found in the higher extent of specialisation of teachers in secondary education, compared to primary education. In primary education, students have only one teacher teaching every subject for an entire year. This situation makes it more straightforward to establish meaningful relations between students and teachers. In contrast, when a student enters secondary education, he/she has different teachers for different subjects. While, during the first two years of secondary education, the number of teachers is still manageable (with one teacher teaching two or three subjects), this division becomes more pronounced once the student enters the higher grades of secondary education

(a different teacher for every subject). With students having so many teachers, and conversely, teachers having a great number of students, it may be difficult to establish social networks between students and teachers, thus contributing to a decline in the perceived student-teacher relationship. However, further research is needed to explore this hypothesis.

Relation between engagement trajectories and moment of dropout

Our second interest was in the relation between engagement trajectories and the moment of dropout. Our results indicate that for both types of engagement, the students in the 'High and Stable' subgroup are less likely to drop out, compared to students in the other subgroup(s). Accordingly, we confirm the results of Archambault et al. (2009) and Janosz et al. (2008). Although students in this 'High and Stable group' were less likely to drop out, there were, nevertheless, a number of these students who did not complete their secondary education and receive a diploma. Accordingly, we extended the previously mentioned studies by not only considering whether these students dropped out, but also when they dropped out. A closer look at the survival curves of the different subgroups revealed some important differences regarding this moment of dropout. The nearly stable progress of the survival curves of the 'High and Stable group' of both engagement constructs, indicates that the majority of the students remain in education until graduating, and if these students do drop out, they are at least in Year 10 before doing so. By the end of their secondary education, more than 95% of the students are still in education and on their way to graduation.

For other subgroups of engagement, a different story emerges. Following Rumberger (1987), our results indicate that dropout can indeed be considered the result of a process of disengagement. These disengaged students begin to leave full-time secondary education very early. For both emotional and behavioural engagement, students in the non-stable or low subgroups begin leaving education in Year 8 and by Year 9, 10% to 15% of the students in these subgroups have already dropped out of school.

With respect to behavioural engagement, both subgroups have nearly the same survival curve in the beginning but starting in Year 11, we observe a switch in

this survival rate with students in the 'High and Decreasing group' having a higher incidence of dropping out. By the end of their secondary education, more than one third of the students in this 'High and Decreasing group' have left secondary education. Although both groups have an increased probability of dropping out, it seems that the instability of behavioural engagement has detrimental effects on the graduation rate of students.

The differences between the subgroups are even higher in the case of emotional engagement. Although both subgroups ('Low group' vs. 'High and Stable group') of emotional engagement demonstrated a decline, one subgroup started at a significantly lower level. It seems that this lower starting level of emotional engagement leads to an increased probability of dropping out. The steep survival curve also indicates that students who report a poor relation with their teachers are quite radical in their decision. When there is a poor student-teacher relation at the start of secondary education, this relation often worsens during secondary education, and thus, these students are likely to turn away from school. Although this corroborates the findings of Croninger and Lee (2001) and Lee and Burkam (2003), it is striking that this dropout mentality exists so early in the educational process.

Background characteristics of dropouts and disengaged students

Concerning the latent class growth estimate as an indicator of the likelihood to dropout, our results confirm previous research. For both engagement constructs (with only slight differences), we conclude that students demonstrating a lower initial cognitive ability, starting in a remedial class, living in a lower SES family, having a minority background and being retained in a grade in secondary education are more likely to leave school prior to obtaining their diploma/certification (see, e.g., Alexander et al., 2001; Lamote, Van Damme, Van Den Noortgate et al., 2013; Lee & Burkam, 2003; Marks, 2007; Rumberger, 2011).

As for the latent class regression estimates predicting class membership in the different engagement subgroups, we confirm the previous research. Male students and students with a low level of achievement who were retained in grade or started in a remedial class were more likely to be member of the 'Low' and/or

'High and Decreasing' subgroups. Students with a low socio-economic background and minority students were also more likely to be members of these subgroups (Archambault et al., 2009; Blondal & Adalbjarnardottir, 2012; Bingham & Okagaki, 2012; Hughes & Kwok, 2007; Van de gaer et al., 2009). Caution is warranted as to not draw any causal conclusions from these findings. For example, previous grade retention seems to be related with a decreasing or an initially low engagement, but we cannot say that grade retention causes this disengagement. It is also possible that grade retention is already the outcome of disengagement. Most likely, the link between grade retention and engagement is part of a cyclic process that can be better explained by the participation-identification model of Finn (1989).

Practical implications

Our results have some practical implications, related to disengagement and to dropping out. Although the group of students at risk for future dropout is characterised by a large heterogeneity, it is clear that the group of disengaged students deserves particular attention. In this study, disengagement was expressed in lower attitudes towards homework, and by more negative relations between students and teachers. As these expressions are often observable, they are also malleable. If a teacher observes signs of disengagement, he or she can look for the possible sources, which can then be a starting point for further intervention (e.g., remedial teacher/class, mentor). In this sense, addressing early school leaving by intervening when a student begins to disengage from education becomes more straightforward compared to intervening on the negative effects of, e.g., the student's SES or gender. Disengagement is a gradual process, and accordingly, it allows teachers to intervene at different moments in time, including the very first signs of the student's disengagement. Furthermore, student engagement can be addressed in the classroom, while intervening on the other predictors of dropout usually occurs at the meso or macro level of the educational system.

An early intervention also seems to be appropriate if one examines the survival curves for dropping out. Because students start leaving full-time education at a very early grade, these students have no diploma or certification from

secondary education, and consequently, their opportunities in the labour market compared to graduates are significantly reduced . Students who drop out already show signs of disengagement when they enter secondary education. Therefore, teachers should be attentive and sensitive to these early differences in engagement, and interventions should begin when the very first signs of disengagement appear.

Limitations and future research

The results of our study should be interpreted within the context of certain limitations.

The first limitation concerns the definition and measurement of school engagement. Although the literature agrees on the multidimensionality of the concept, Reschly and Christenson (2012) note the different conceptualisations that are still prevalent. While most studies recognise that engagement consists of an affective component and a behavioural component, several studies add other components to this engagement (see e.g. Fredricks et al. 2004; Kortering & Christenson, 2009). The inconsistency in the number of components inevitably leads to different measurements of the components of student engagement. Therefore, consistent with Fredricks et al. (2004), we selected scales that we believed best represented the conceptualisation of engagement. However, this was only possible for two dimensions as we were not able to find a feasible indicator for cognitive engagement. Although some scales seemed to be good indicators of cognitive engagement, an examination of the items that composed these scales revealed that most of the items assessed a behavioural aspect rather than a psychological component. Accordingly, the items serve as a potential indicator of behavioural engagement rather than cognitive engagement. Additionally, we are aware that some indicators of engagement components that we used, are debatable.

A second limitation concerns the method used to assess engagement. As Fredricks and McColsky (2012) noted, most measures of engagement are based on self-report surveys. Most of the time, these surveys only assess general engagement components, and even though the use of such self-report surveys makes it possible to assess a large number of students, there is a possibility that the answers do not reflect the behaviours or relations that we seek to assess. For example, the items on

self-report surveys usually consist of very broad wording (in our case, e.g., "When I want to do something nice, I still complete my homework first"). Such broad statements are insensitive for differences in engagement related to the specific learning domain.

Third, as clearly indicated by the differences between the intercepts of the 'High and Stable group' and the 'Low (and Decreasing) group' for both engagement constructs, students enter secondary education with marked differences in engagement. However, from our data, we are not able to test for possible sources of these differences, as we have no information on (the development of) student engagement in primary education. It is possible that these differences were already present in primary education and that the transition from primary to secondary education had a significant effect on the engagement level of the student. Accordingly, future research should focus on data spanning the entire educational career of the student, starting in primary education with measurements of student engagement through the end of the student's secondary education. If this is not feasible, we suggest research that focuses on the transition from primary to secondary education.

In this study, our interest was mainly in student characteristics related to engagement and dropout. Although in our analysis we clustered students within schools, we did not examine the effects of school characteristics on the development of engagement and dropout. However, previous research on these topics suggests differences between schools for both engagement (Fullarton, 2002) and dropout (Lamote, Van Damme, Van Den Noortgate et al., 2013). These differences between schools can be a focus for future research on engagement and dropout.

CONCLUSION

From our results, we conclude that, for behavioural engagement, students with unstable pathways encounter increased hazards that lead to dropping out, compared to students with stable, normative patterns. For emotional engagement, it seems that the greater hazards associated with dropping out were mainly the result of the low level of engagement at the beginning of secondary education and that the development of engagement had less effects on the likelihood of dropping out. For both engagement dimensions, these increased hazards are already present in the early years of secondary education.

Chapter 4

Is the cure worse than the disease?

A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept.

Based on: Lamote, C., Pinxten, M., Van Den Noortgate, W., & Van Damme, J. (2013). Is the cure worse than the disease? A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept. *Manuscript submitted for publication.*

ABSTRACT

Holding students back when they do not meet a specific attainment level is common practice in a lot of countries. However, this practice is not without controversy and recent studies point at the negative effects of grade retention, especially in the long-term. The majority of these studies focussed on grade retention in primary education. In our study, we focussed on the effect of grade retention in Grade 8 on language achievement and academic self-concept. We matched students who were and were not retained based on their propensity to be retained and compared both groups using a growth curve analysis. The basic treatment 'grade retention vs. promotion' was extended with the certificate these students received at the end of Grade 8. We can conclude that grade retention has a negative effect on the achievement of retained students in the long run, but has no negative effect on academic self-concept. When we take into account the certificate they received, it seems that following the teacher's advice to change track is a better decision than repeating the grade in the same track.

INTRODUCTION

At the end of each school year, the majority of students can start a long summer holiday with the prospect of starting the next school year in a higher grade. Unfortunately, not every student has such a bright outlook. Yearly, a substantial number of students repeat a grade in primary or secondary education. Based on PISA 2009 data, an international average of 13% of 15-years old students repeated one or more years in primary or (lower) secondary education. These students, grade retainees or repeaters – and more specifically repeaters in lower secondary education, – are the focus of the present study.

The non-promotion of students is not without controversy, given that a considerable amount of research on grade retention points at the negative effects of this non-promotion, both for the student and for society at large. However, in their meta-analysis, Allen et al. (2009) concluded that the results of the studies under consideration were strongly related to the design quality of the examined studies. Studies reporting the strongest negative effects of grade retention often had the weakest methodological ground, whereas studies with a sound methodology reported only a small negative or no effect of grade retention. However, 'no effect of grade retention' is no grist to the mill of the proponents of grade retention, because if there is no effect then what is the benefit of grade retention?

Research on grade retention already has a long history. As Jackson (1975) noted in his review: research on grade retention dates back to the beginning of the previous century. Over the years, this research was mainly conducted at the primary school level and focused on differences in achievement and non-cognitive development between retained and regularly promoted students during their school career (e.g. Bonvin et al., 2008; Hong & Yu, 2008; Wu, West, & Hughes, 2010).

Yamamoto and Byrnes (1987) concluded that older students perceived grade retention as more stressful compared to younger students, but despite this, none of the studies that were labelled as methodologically sound by Allen et al. (2009) considered the achievement or non-cognitive outcomes of grade retention in secondary school. To date, only few studies that examined the effect of retention in secondary education on these outcomes would meet the strict requirements of Allen

et al. (2009) (see e.g. Ehmke et al., 2010; Uysal, 2010). In the current study, we will focus on grade retention in secondary education in Flanders, Belgium.

In the following paragraphs, we present a concise overview of the current knowledge on the effect of grade retention (both in primary as in secondary education) and we give an overview of the research questions in this study. After the theoretical paragraph, we describe our sample and explain our methodological approach. After this, we present our results and relate them to the current knowledge in the discussion section.

THEORETICAL FRAMEWORK

In an educational system where the grouping of students is age-based, one way of handling the heterogeneity in academic achievement is holding back those students who do not meet the required attainment level of a certain grade. The underlying assumption is that the extra granted year enables them to catch up with the material they did not master. However, the debate on the effectiveness of this practice is as old as the practice itself. In 1975, Jackson reviewed 44 studies and concluded that there is not enough evidence that grade retention is more beneficial for students with academic difficulties than grade promotion. This finding was more or less confirmed in subsequent reviews of Holmes (1989) and Jimerson (2001), especially regarding the long-term effects of grade retention. In the most recent meta-analysis by Allen et al. (2009), this non-effectiveness of grade retention was toned down because most of the studies reporting strong negative outcomes of grade retention, were studies with (serious) methodological flaws. However, Allen et al. only considered studies conducted in North America and published before June 2007. More recently, a number of studies were published in which the authors took into account the methodological concerns. Still, these studies mainly focused on grade retention at the primary school level. The results of these studies can be divided into two broad categories, based on the considered outcomes: academic achievement on the one hand, psychosocial functioning on the other.

Previous research on the effects of grade retention on academic achievement yielded different results depending on whether short or long-term effects were considered. In the short term, during the year of retention, mainly positive effects were found regarding academic achievement: at the beginning of the retention year, non-promoted students outperformed their new (younger) classmates (Alexander et al., 2003; Karweit, 1999; Wu et al., 2008). However, throughout the repeating year, the achievement of the non-promoted students declines, and by the end of the repeating school year they are behind of their classmates (Alexander et al., 2003; Bonvin et al., 2008; Wu et al., 2008). In general, studies considering the long-term effects of grade retention conclude that the positive effect on cognitive outcomes in the (beginning of) the repeated year usually diminishes (Bonvin et al. 2008; Karweit, 1999), disappears (Jimerson et al., 1997) or even shifts to a negative effect (Alexander et al., 2003). This long-term null or negative effect is one of the main arguments for the opponents of grade retention.

Regarding the psychosocial outcome of grade retention, two theories are commonly cited: the social comparison theory of Festinger (1954) on the one hand and the labelling theory (Becker, 1963) on the other. The social comparison theory is most often used by teachers and parents in support for grade retention. When repeating a grade because of low achievement, a student will be exposed to the subject matter of that particular grade for a second time, but will find himself in a classroom with students who never studied this subject matter. Because of this (assumed) head start, the retained student will feel more competent in this matter compared to his new classmates. As a result, this head start provides the retained student with a higher level of perceived competence and a higher level of academic interest (Hong & Yu, 2008). It is assumed that the retained student will regain confidence because he can keep up again. However, empirical results indicate that this positive effect seems to be of short duration: by the end of the school year, this positive effect disappeared (Bonvin et al., 2008). In contrast with the positive effects assumed by Festinger's social comparison theory, Becker's (1963) labelling theory postulates that students retained in grade are given the label of 'stupid' or 'failures' by their new, younger, classmates. This label of 'repeater' can potentially lead to a decline in school engagement and self-esteem. The effect of grade

retention on the social development of students seems to be related with the moment of grade retention: early grade retention (i.e. at the beginning of primary education) seems to have a smaller negative impact on a student's self-concept, compared to later grade retention and as such, students retained in kindergarten do not suffer from the 'retainee' label as much as students retained in secondary education (Shepard, 1989).

Secondary education and grade retention in Flanders

In order to fully grasp grade retention practice in secondary education in Flanders, it is necessary to describe the ways to grade retention in the Flemish educational system[x]. At the end of each school year, every student not only receives a school report, but also a certificate stating in which grade/track this student is allowed to start the next year. Three different certificates can be issued: an 'a-certificate' (student is allowed to proceed to the next year, without restrictions for certain tracks), a 'b-certificate' (the student has two options: repeat the current grade in the same track, or go to the next year but with restrictions for certain tracks) or a 'c-certificate' (the student must repeat the current grade). For students who receive a 'b-certificate' in a tracked educational system, repeating a grade is sometimes more attractive than enrolling in a lower track (academic track vs. technical/vocational track), because of the negative image of the lower tracks (especially the vocational track). This grade retention on a 'voluntary' basis can have a different effect on the observed outcomes, because students are more likely to perceive this grade retention as a 'second chance' instead of real 'failure' (Kloosterman & de Graaf, 2010).

Same-grade or same-age?

Before turning to the research questions, it is important to highlight one specific aspect of research on grade retention. When examining grade retention, a researcher can choose between two common approaches: same-grade comparisons or same-age comparisons. A same-grade comparison compares students who were retained in grade with students in the same grade (so with a different age). This comparison is usually in favour of the retained students in the year of retention, (both for achievement as for non-cognitive outcomes) but the positive effect decreases over

time (Allen et al., 2009). The second approach is the same-age approach in which grade retainees are compared with promoted students of the same age, but one year higher. This approach usually indicates more negative effects of grade retention in the year of retention, but these negative effects become more positive over time (Allen et al., 2009). A visual representation of these approaches can be found in Figure 7.

Research questions

We will focus on two outcomes in the current study: language achievement and academic self-concept and we will examine the effect of grade retention on the outcomes in the short-term (year of retention) and in the long-term (years after retention). As noted in the literature section, taking into account short- as well as long-term effects is vital in grade retention research.

First, we will examine the academic growth of promoted vs. retained students with a focus on the (Dutch) language achievement. Second, following the results of Hong and Yu (2008), we are interested in the development of the academic self-concept of students retained in grade vs. regularly promoted students. For both outcomes, we also estimate the effect of grade retention after receiving a 'b-certificate' (voluntary) or 'c-certificate' (obliged) on both outcome variables. Summarizing, four research questions are at stake:

1. Is there a difference between retained and promoted students in psycho-social growth after grade retention, until the end of secondary education?
2. Is there a difference between retained and promoted students in academic growth after grade retention, until the end of secondary education?
3. Is there a difference in psycho-social growth for students retained in grade (or not) after receiving a 'b-certificate', compared to the other groups of students ?
4. Is there a difference in academic growth for students retained in grade in grade (or not) after receiving a 'b-certificate', compared to the other groups of students?

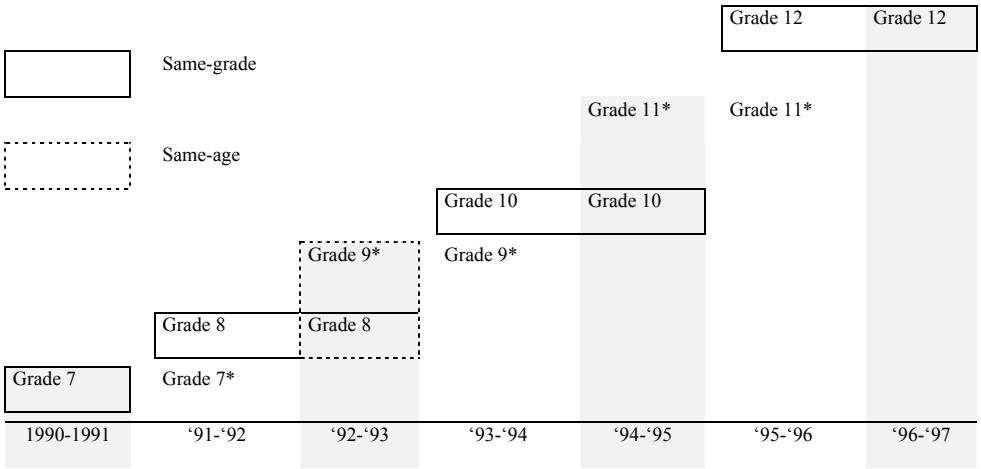
METHOD

Data

For this study, the data were drawn from the Flemish longitudinal 'LOSO'-project (Van Damme et al., 2002). This research project initiated in 1990 and followed a cohort of 6,411 students through secondary education, with a follow-up in higher education or the first years on the labour market. This database contains information on student background characteristics, teacher/class characteristics and school characteristics. Most of the background characteristics were measured once (at the beginning of secondary education), except for the achievement related and non-cognitive measures. These variables were measured in several grades.

From the original sample, we selected a subsample of 3,900 students in 50 schools. These students were enrolled in the Grade 7 in 1990-1991 and in Grade 8 in 1991-1992. More specifically, we selected all students who proceeded through secondary school without a delay (n = 3,616) and students who were retained in grade for the first time in Grade 8 (n = 284). This resulted in a more restricted sample, but this was necessary to obtain three measurement occasions, which is needed for examining a nonlinear growth curve (Singer & Willett, 2003).

Due to data restrictions (see Figure 7), our results are based on a same-grade comparison.



* no data available. For students not retained in grade, data on achievement and academic self-concept is available in Grade 7, 8, 9 and 10. For retained students, data only available in Grade 8 ('92-'93), Grade 10 ('94-'95) and Grade 12 ('96-'97).

Figure 7: same-age vs. same-grade and available data

Achievement

In the present study, we use the Dutch language achievement of the students as an indicator of academic growth. This language achievement was tested several times during secondary education. For regularly promoted students, the test was administered at the beginning of Grade 7 (1990-1991), end of Grade 7 ('90-'91), 8 ('91-'92), 10 ('93-'94) and 12 ('95-'96). Retained students administered the Dutch language achievement test also at the beginning and end of Grade 7, at the end of Grade 8, a second time (but one year later) at the end of Grade 8, and finally also in Grade 10 and 12. The tests were developed as part of the LOSO-project and are composed of curriculum-relevant multiple choice items. Different tests versions were administered in different grades and tracks, with adopted difficulty levels. These test versions shared overlapping items and by means of item response theory (IRT), calibrated test scores were computed. The internal consistencies for the different tests ranged from $\alpha = .66$ to $\alpha = .93$, indicating moderate to very high internal consistency.

Academic self-concept

The academic self-concept scale of the LOSO-project stems from a general well-being questionnaire, which was administered four times (end of Grade 7, 8, 10, 12) for regularly promoted students and four times for grade retainers (end of Grade 7, 8, 10, 12). This scale was composed of 9 items, referring to the general academic self-concept of the student. Sample item: 'I think that I am good at learning' (For a complete overview, see Van Damme et al., 2002). The internal consistency of this scale varied over the years from $\alpha = .78$ to $\alpha = .81$, which can be considered as a satisfactory internal consistency.

Matching

To analyse the effect of grade retention, we could not simply compare the scores of repeaters with the score of non-repeaters, because of the differences between the two groups on multiple background characteristics, such as achievement, gender, intelligence. Students can also not regarded as randomly assigned to a treatment condition (grade retention) and a control condition (promotion). Therefore, a matching of the students based on all the variables related to the treatment seemed

the most appropriate way to investigate the effect of grade retention. However, this set of pre-treatment variables can become very large, and consequently, matching on all these variables becomes nearly impossible. A solution for this, is the use of propensity score (PS) matching (Rosenbaum & Rubin, 1983). A PS matching approach allows us to compare groups of students who are implicitly different, by controlling for several variables related to the treatment. The propensity score, which is a function of covariates, refers to the conditional probability that a unit with vector x of observed covariates will be assigned to treatment condition 1/0 (Rosenbaum & Rubin, 1983). Instead of using all pre-treatment variables in computing the propensity score, we follow the approach of Imbens and Rubin (2012) by only selecting the most relevant covariates. Therefore, we included covariates (and interactions/higher order terms) into a logistic regression model one at the time and evaluated the model fit. We repeated this until adding another covariate did not improve the model fit. With this model, the PS was computed. In a next step, this PS is used to match students with a similar PS, constructing homogeneous treatment and comparison groups. In this step, we opted for a full matching approach (Stuart & Green, 2008) instead of the more common '*k*:1 nearest neighbour matching' or 'subclassification/stratification' approach. As Stuart and Green note, the full matching can be thought of as a compromise between 'nearest neighbour matching' and 'subclassification'. In 'nearest neighbour matching', one matches every treated individual to k comparison units. As Thoemmes and Kim (2011) noted, this kind of matching is popular in research using a PS-approach, but using this *k*:1-matching can result in a very small dataset, because many comparison units are discarded. This, in turn, will lead to biased estimates. Another popular approach for reducing imbalance on the covariates between treated and control group, is subclassification. With subclassification, students with a similar propensity score are grouped together in subclasses. However, it is not always clear how many subclasses are needed to achieve balance on the covariates. Following the advice of Rosenbaum and Rubin (1984), five or six subclasses are sufficient, but especially in large samples, it is not always clear how many subclasses are needed to achieve balance on the covariates (or it is very time-consuming to achieve this balance).

The full matching automatically determines the number of subclasses. With full matching, matched sets of treated (retained students) and control (not retained in grade) students are formed and each matched set contains at least 1 grade retainee and at least 1 student who was not retained in grade. In every subclass, students received a weight with students retained in grade receiving a weight of '1' and the matched students (promoted students) receiving a weight equal to the retained students in the subgroup, divided by the number of non-retained students in the subgroup (for more information on this weighting, see: Stuart and Green (2008) and the documentation of the MatchIt software package (Ho, Imai, King & Stuart, 2011)). We evaluated the balance of the matched groups by looking at the standardised bias. This standardised bias is defined for every covariate and is calculated as the difference in means between the treatment and control group, divided by the standard deviation of the full matching group (treatment and control group). Standardised biases smaller than .25 indicate a good match (Ho, Imai, King, & Stuart, 2007). The full matching was done by using the MatchIt-package in R 2.15.1 (Ho et al., 2011; R Core Team, 2012).

Analysis

After matching on the PS, we use the weights resulting from the full matching and include these weights in a growth curve analysis, by making use of the 'lme4-package' in R 2.15.1 (Bates, Maechler, & Bolker, 2012). In the subsequent growth analyses, we account for the multilevel nature of the data with measures at Level 1, students at Level 2 and schools at Level 3 (Singer & Willett, 2003). Time was coded as '0' for Grade 8, '1' for Grade 10 and '2' for Grade 12 (0, 1, 4 for Time²).

RESULTS

Full matching

Following the approach of Imbens and Rubin (2012), the estimation of the PS was based on 32 covariates and interactions (see Appendix 2 for a full list of covariates used in the final PS-model). With this set of covariates, all imbalances between matched subgroups were removed as no standardised bias was higher than .25.

Based on the estimated PS, we were able to match 215 grade retainees (75.7% of original sample of grade retainees) with 1,380 promoted students (38.1% of original sample of promoted students). The matching and overlap between grade retainees and promoted students, together with the weights received, is illustrated in Figure 8. The size of each circle reflects the weight of that student, a larger circle indicates a higher weight.

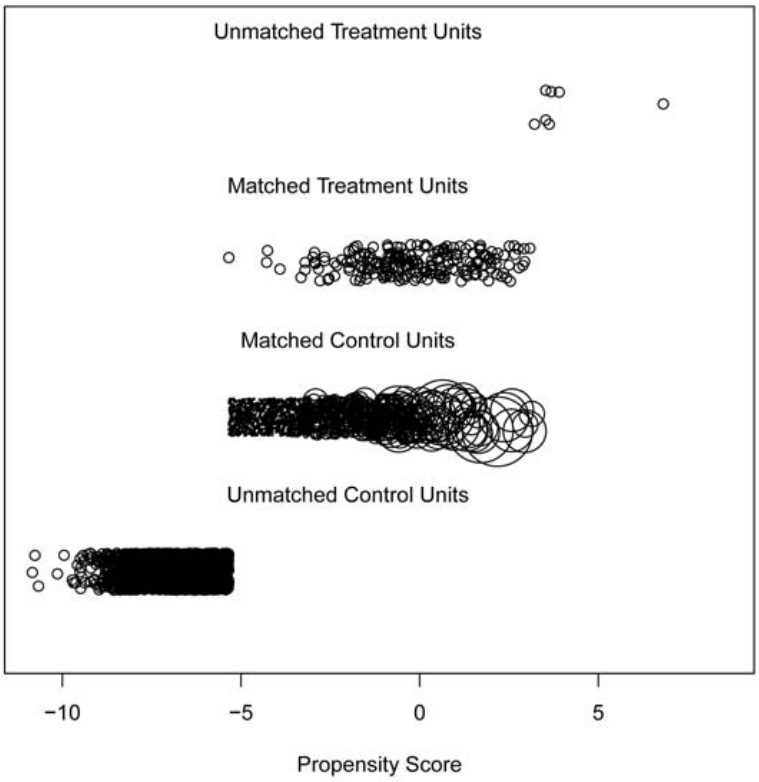


Figure 8: distribution of propensity scores

In the following paragraphs, we will focus on the effect of grade retention on academic achievement and academic self-concept separately. When discussing these effects, we first focus on the overall effect of grade retention, followed by the results of analysing the effect of repeating of groups with different certificates.

Table 13: Parameter estimates

Parameter		Language achievement				Academic self-concept			
Fixed effects		Est.	*SE*	Est.	*SE*	Est.	*SE*	Est.	*SE*
Intercept	β_0	-0.895***	*0.072*	-0.663***	*0.079*	3.041***	*0.023*	3.211***	*0.032*
Grade retention	β_1	-0.084	*0.057*			0.455***	*0.042*		
Time	β_2	-0.037	*0.040*	-0.676***	*0.074*	0.688***	*0.029*	0.428***	*0.053*
Time²	β_3	0.001	*0.020*	0.287***	*0.036*	-0.219***	*0.014*	-0.151***	*0.026*
Grade retention * Time	β_4	0.628***	*0.130*			-0.877***	*0.092*		
Grade retention * Time²	β_5	-0.534***	*0.066*			0.300***	*0.047*		
Grade retention after 'b'	β_6			-0.257***	*0.080*			0.303***	*0.060*
Grade retention after 'c'	β_7			-0.388***	*0.083*			0.271***	*0.059*
Promoted after 'b'	β_8			-0.365***	*0.056*			-0.260***	*0.041*
Grade retention after 'b' * Time	β_9			1.239***	*0.171*			-0.778***	*0.124*
Grade retention after 'c' * Time	β_{10}			1.266***	*0.210*			-0.377***	*0.144*
Promoted after 'b' * Time	β_{11}			0.892***	*0.087*			0.352***	*0.062*
Grade retention after 'b' * Time²	β_{12}			-0.778***	*0.086*			0.318***	*0.063*
Grade retention after 'c' * Time²	β_{13}			-0.867***	*0.108*			0.103	*0.075*
Promoted after 'b' * Time²	β_{14}			-0.398***	*0.043*			-0.085**	*0.031*
Random effects		Est.	*SD*	Est.	*SD*	Est.	*SD*	Est.	*SD*
School		0.213	*0.462*	0.211	*0.459*	0.005	*0.071*	0.006	*0.078*
Student		0.175	*0.418*	0.170	*0.413*	0.117	*0.342*	0.116	*0.341*
Residual		0.300	*0.548*	0.285	*0.534*	0.144	*0.380*	0.138	*0.373*
Deviance		7,469		7,342		4,316		4,187	

Academic achievement

Overall effect

When we compare the achievement of grade retainees with the achievement of promoted students which are equally at-risk to be retained in grade (see Table 13), these grade retainees performed the same as promoted students a year before in Grade 8 (β_1 = -0.084, SE = 0.057, p = .140).

In the long run, promoted students perform at a constant level. Retained students, on the other hand, experienced strong negative effects on the long run. These effects become more clear in Figure 9.

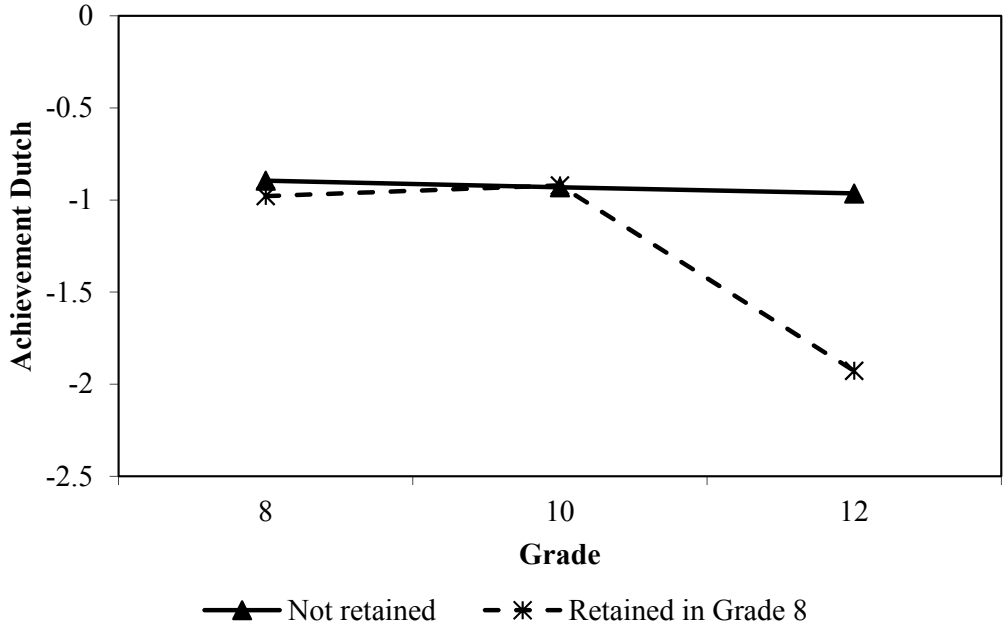


Figure 9: overall effect of grade retention on achievement

Grade retention after b- or c-certificate

Instead of treating the group of grade retainees as one group, we make a difference between grade retention after receiving a 'c-certificate' (obliged) or a 'b-certificate (voluntary). Based on the results reported in Table 13, it is clear that students who choose themselves to repeat the year and students who received a 'c-certificate',

achieve significantly lower in the retention year compared to promoted students equally at-risk for grade retention a year earlier. Students who received a 'b-certificate' at the end Grade 8 but chose to change over to the Grade 9 in another track, already achieved lower in Grade 8 than the regularly promoted students with an equal risk to be retained. As illustrated in Figure 10, there are also differences between the subgroups in the long run. Both subgroups of retained students showed a fairly stable achievement on the short term but a significant strong decline in the long run. At the end of secondary education, both groups of retained students had a significantly lower achievement (at $p < .001$ level, not shown in the table) compared to the two groups of promoted students.

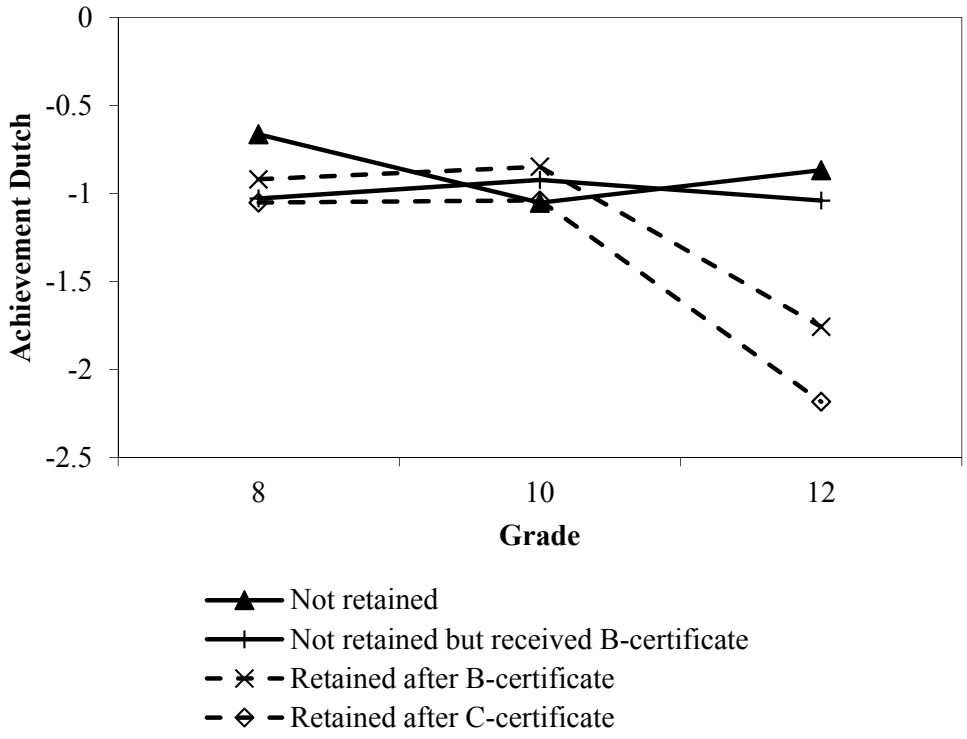


Figure 10: differential effects of certificates on achievement

Academic self-concept

Overall effect

If we focus on the effect of grade retention in Grade 8, it is clear that students retained in the second year of secondary education have a significantly higher academic self-concept in the year of retention (β_1 = 0.455, SE = 0.042, p < .001). This positive effect of grade retention is compared to the younger grade-mates who had an equal risk to be retained in grade, but who were promoted instead. However, in the long run, this advantage disappears and in Grade 12, there was no significant difference between both groups (effect of grade retention in Grade 12 (not shown in Table 13): β = -0.098, SE = 0.059, p = .096). This becomes more clear when we plot these growth curves in Figure 11.

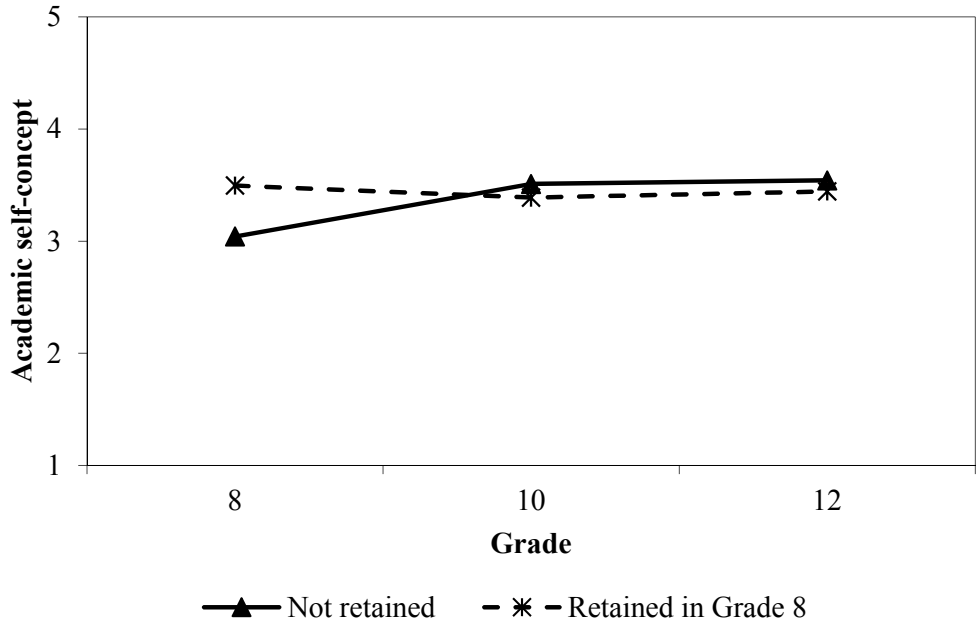


Figure 11: overall effect of grade retention on academic self-concept

Grade retention after b- or c-certificate

When we focus on the certificate the students received at the end of the first time in Grade 8, we find a difference on the overall development of academic self-concept,

especially for the group of promoted students. In this group, as Figure 12 indicates, there is a difference regarding the intercepts of students promoted after receiving a 'b-certificate' or students promoted after an 'a-certificate'. Students who received a 'b-certificate' at the end of Grade 8 had a significantly lower academic self-concept (β_8= -0.260, SE = 0.041, $p < .001$), compared to the promoted students with an 'a-certificate'. However, as they proceed through secondary school, they show a positive growth and even surpass the academic self-concept of the 'a-certificate'-group in the 12th grade (effect of grade retention in Grade 12 –not shown in the table: $\beta = .103$, SE = 0.044, $p = .017$). As for the two groups of retained students: both groups still have a higher academic self-concept the second time in Grade 8, compared to the promoted group. But again, this advantage diminishes through secondary education and at the end of secondary education, there is no difference between these grade retainees and students receiving an 'a-certificate'.

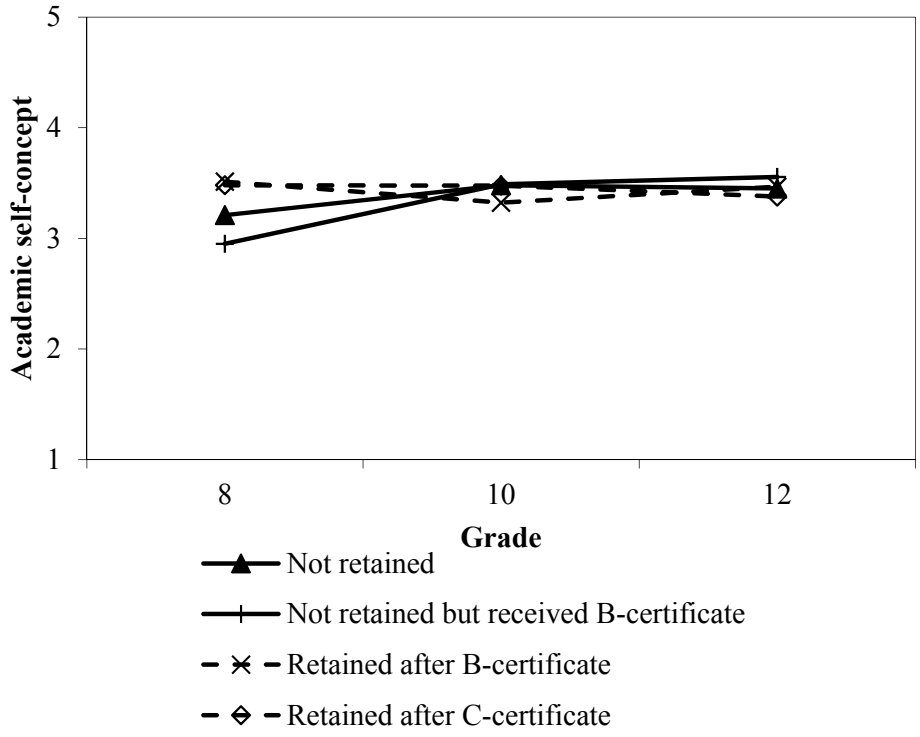


Figure 12: differential effect of certificates on academic self-concept

DISCUSSION

In this study, we examined the effect of grade retention in Grade 8 on language achievement and academic self-concept. Overall, our results show that retaining students in grade in some cases had a positive effect in the short term, but this positive effect disappeared in the long term. When we split up the treatment and control groups into the certificate the student received, it became clear that these group differed in long-term effects.

With this study, we meet several limitations of previous studies. First, instead of focussing on the effect of grade retention at the primary school level, our interest was mainly in the consequences of grade retention in secondary education. We did not limit our study to one outcome, but considered the achievement as well as a psychosocial outcome of the student, both in the short term and in the long term. With the long-term focus, we extended the study of Ehmke et al. (2010). Second, in order to meet the critiques on the methodology of research on grade retention, we used a propensity score (PS) matching approach. PS-matching allowed us to compare retained students with a group of students with the same propensity to be retained in grade, but were promoted instead. We found a good match between 215 grade retainees and 1,380 promoted students. The remaining promoted students had a PS which was not comparable with the PS of students in the treatment group and these students were discarded from further analyses. After the matching, we first examined the overall effect of grade retention, followed by a deeper examination of the growth in academic achievement and academic self-concept after promotion and after receiving an b-, or c-certificate (i.e. after voluntary or obligatory grade retention).

Our results of the overall effect of grade retention in the short term (the year of retention) indicated that repeating a grade positively affected students' self-concept but did not lead to increased performance. In line with Ehmke et al. (2010) we can conclude that grade retainees do not perform better compared to a comparable group of students the year before, despite the fact that these grade retainees attend Grade 8 for the second time and are one year older. However, as the significant effect on academic self-concept indicated, these grade retainees felt

more competent in the year of retention. In line with Hong and Yu (2008), this provides support for the social comparison theory of Festinger (1954) because in the new classroom, grade retainees will have 'a head start' as they receive the subject matter for the second time. As a consequence, grade retainees are more likely to feel more competent compared to their new classmates.

Taken together, we can conclude that grade retention does not harm the student's achievement and gives a boost to the grade retainee's academic self-concept. This picture becomes more negative when effects in the long-run were considered. After the year of retention, especially the achievement of the grade retainees decreased. A possible explanation for this (sudden) decrease is in the tracking: in the group of grade retainees, only a minority of the students (14.4%) are in the academic track by the end of Grade 12. Especially in the vocational track, language education has a less important place in the curriculum.

This finding further extends the results of previous studies that investigated the effects of grade retention in primary education (Alexander et al., 2003; Bonvin et al., 2008; Wu et al., 2008). It seems that grade retention only leads to a higher self-concept and rather good achievement in the repeating year.

As a specific aspect of the tracked Flemish educational system, we were interested in the effect of the certificate received. The effects of this certificate on academic self-concept are especially interesting in the long-run, and more specifically for the students who received a 'b-certificate'. Our results show that students who ignore the teacher's advice to enrol a lower track only benefits from this in the short-run. In the long run, it are the students who decided to change track, that show a positive growth. Changing to a lower track when the teacher gives the advice to do so, seems to be a good decision. Partially in line with Wouters, De Fraine, Colpin, Van Damme and Verschueren (2012), promoted students changing over to another track experienced an increase in academic self-concept. However, Wouters et al. found that this increase was of short duration, while in our case this increase lasts until the end of secondary education. It should be noted that the sample of Wouters et al. was more general compared to the sample in our study, which consisted of students who were all at risk for grade retention.

Moreover, since we only focussed on changing tracks in Grade 8, it is possible that students who changed track (again) after Grade 8 give this growth curve a positive 'boost'.

A similar conclusion holds true for the effect of receiving a b- or c-certificate on academic achievement. In the year of retention, students who repeated the year on voluntary basis achieved the same as promoted students. However, in line with Alexander et al. (2003) the effect of grade retention for both groups of grade retainees shifts to a negative effect on the long-term.

This conclusion, in combination with the results of the effect of grade retention on academic achievement, seems to be the most important conclusion of the effect of grade retention. When comparing groups of students with the same propensity to be retained in grade, it becomes clear that grade retention is not the best solution for low-performing students and that 'the cure' is indeed worse than 'the disease'. After splitting up the students on their obtained certificate, it seems that a good advice on future tracks given by the teacher is more effective than repeating a grade.

Limitations and suggestions for future research

The results of our study should be interpreted within the context of some limitations. The first limitation is in the specification of the propensity score. Although students receiving a 'c-certificate' and a group of students receiving a 'b-certificate' all experience the same treatment (grade retention), this case can also be thought of as students receiving a 'multiple treatment' (Imbens, 2000) (with all the relating difficulties and problems of the 'multiple treatment' approach (Stuart, 2010)). However, in our case, we looked at the treatment as a 0/1 case and looked at the differential effects of the certificates on later achievement/academic self-concept on a descriptive level. A second limitation follows from data restrictions, and is also a suggestion for further research. As Yamamoto and Byrnes (1987) noted, grade retention is perceived as more stressful by older students, compared to younger students. Therefore, it seems necessary to focus on 'late' grade retention (Grade 9-12) and relate this to academic self-concept, achievement but also to the decision of early school leaving (Jimerson et al., 2001).

Chapter 5
When you have beaten the odds: success of grade retainees in higher education

Based on: Lamote, C., Van Den Noortgate, W., & Van Damme, J. (2013). When you have beaten the odds: success of grade retainees in higher education. *Manuscript submitted for publication.*

ABSTRACT

Several studies point at the negative (or non-positive) effects of grade retention on different outcomes (e.g. Wu, West & Hughes, 2008). Some researchers related grade retention to outcomes on the post-secondary education level, and concluded that students retained in primary and/or secondary school have a lower chance to enrol in higher education and have little success in higher education (Fine & Davis, 2003; Ou & Reynolds, 2010, Pustjens et al., 2004). However, there were some methodological issues with the previous studies, which we will cover in this study. In the current study, we focus on the effect of grade retention in secondary education, on the enrolment in higher education and the success (in the short- and long-term) in higher education. By using data from a longitudinal study, and by making use of a propensity score matching approach, we can draw stronger conclusions on the effect of grade retention on the outcomes we mentioned. Based on this matching, we found that students retained in grade in secondary education, participate less in higher education. Once they are enrolled, we found no negative effect of retention in the early years of secondary school on success in higher education. In contrast, retention in the last three years of secondary school seems to have strong negative effects on the success in higher education, both in the short-term as well as in the long-term.

INTRODUCTION

In 2010, two years after the break-out of the financial crisis, the President of the European Commission declared in the preface of the EU2020 framework that 'EU 2020' is "…about more jobs and better lives. It shows how Europe has the capability to deliver smart, sustainable and inclusive growth, to find the path to create new jobs and to offer a sense of direction to our societies" (European Commission, 2010, p. 2). In order to deliver this smart, sustainable and inclusive growth, the EU sets out 5 headline targets dealing with contemporary problems or situations. One of these 5 headline targets deals with the first part of the growth objectives: Europe needs to deliver a smart growth in order to develop an economy based on knowledge and innovation. To become a knowledge economy, Europe aims at a reduction of the early school leaving rate below 10% and minimum 40% of the 30-34 years old completing tertiary education. In the present paper, we will focus on this 40% target (for a state of the art report on early school leaving, we refer to the most recent book of Rumberger, 2011). At the moment, the EU-rate of graduates in tertiary education is at 35.5% (EUROSTAT, 2013), an increase of more than 10% compared to ten years ago. However, in order to achieve the ambitious EU target of 40%, there's still a lot of work to be done.

Although several student characteristics predict enrolment in higher education, we will focus in this study on one specific predictor of enrolment: the practice of grade retention. Holding students back is a controversial issue because of the potential negative outcomes of this practice. Several studies point at a higher chance to dropout (see e.g. Jimerson et al., 2002) and to lower levels of achievement (Wu et al., 2008) for grade retainees. Grade retention has also negative effects on long-term (educational) outcomes such as enrolment in higher education (Fine & Davis, 2003; Ou & Reynolds, 2010, Pustjens et al., 2004). The difference in enrolment rates in higher education between grade retainees and continuously promoted students was the topic of some studies, but only few studies considered another important aspect of participation in higher education: graduation. This sole focus on mere attendance is somewhat surprising, because, as Hanushek and Woessmann clearly state, "it is the learning and not the attendance that must have the highest priority" (Hanushek & Woessmann, 2012, p. 75).

In the following paragraphs, we give a concise overview of the results of previous research on grade retention. This overview first focuses on the effects of grade retention on overall achievement and non-cognitive outcomes, but afterwards, we give special attention to the studies of Ou & Reynolds (2010) and Pustjens et al. (2004) because of their specific outcome and/or methodological approach. We conclude the literature review with the research questions at stake, followed by a description of the sample and method used in this study. We end the paper by summarizing the results and relating these results to the current literature.

LITERATURE REVIEW

Grade retention, achievement and non-cognitive outcomes

When discussing the effects of grade retention on achievement and non-cognitive outcomes, one should bear in mind that most of the studies were conducted at the primary school level (e.g. Hong & Raudenbush, 2005; Wu et al., 2008). When the secondary school level was considered, the focus is often on the effect of grade retention during primary school on outcomes in secondary school (e.g. Jimerson & Ferguson, 2007), although Yamamoto and Byrnes (1987) noted that being hold back in grade at an older age could lead to more severe consequences. Bearing in mind these limitations, we discuss some of the most important results of grade retention, stemming from some 'landmark' studies.

Holding a student back at the primary school level, does not have the beneficial effects (usually) assumed by teachers and parents. There is an overwhelming number of studies reporting negative or non-effects of grade retention on long-term achievement (e.g. Hong & Raudenbush, 2005; Moser, West, Hughes, 2012, Wu et al., 2008). Although retained students are (most of the time) able to catch up in the year of retention, this advantage is found to diminish or disappear over the years. Recently, Allen et al. (2009) conducted a large meta-analysis of 22 studies which focussed on the effects of grade retention on achievement. They concluded that the low quality of the methodological design was in a lot of studies the cause of the strong negative effects of grade retention. Indeed,

the methodologically sound studies often found no, or only a very small (often non-significant) effect of grade retention. Allen et al. therefore concludes that "given the expense of grade retention (…), a finding of 'no significant difference' for retention on achievement calls into question the educational benefits of grade retention policies" (Allen et al., 2009, p. 493).

Regarding the effects of grade retention on non-cognitive outcomes, previous studies found no effect or a small positive effect of grade retention, depending on the type of outcome and the time-frame. For instance, Wu et al. (2010) found that both in the short-term as well as in the long-term, grade retainees reported a higher behavioural engagement, a higher academic self-concept and a lower hyperactivity. On the other hand, Bonvin et al. (2008) found that the initial positive effects on some non-cognitive outcomes, are levelled out by the end of the retained grade.

Grade retention and higher education

Previous research consistently found that students retained in grade are more likely to drop out of secondary education, compared to similarly low-achieving but not-retained students (e.g. Jimerson et al., 2002). Yet, grade retention does not always lead to dropout since a proportion of students retained in grade successfully graduate from secondary education. Ferguson, Jimerson and Dalton (2001) referred to these students as 'successful failures' since they have 'beaten the odds' to dropout despite the fact they were retained in grade (Fine & Davis, 2003). However, these 'successful failures' were rarely the subject of research, since only a few studies focused on the effects of grade retention beyond secondary education (Fine & Davis, 2003; Jimerson, 1999; Ou & Reynolds, 2010; Pustjens et al., 2004). Jimerson (1999) indicated that students who were retained in grade in primary school, were less likely to enrol in higher education, compared to low achieving but promoted students. When these retained students start in higher education, they show less ambition in their educational career choice (Pustjens et al., 2004) and are less likely to obtain a diploma of higher education. The effect of grade retention on enrolment in higher education seems to depend on the timing of this retention. Ou and Reynolds (2010) found different effects of enrolment for early or late retention

groups. Students who were retained during grade 4-8 were more likely to drop out and less likely to start higher education, compared to students retained in grade 1-3. Fine and Davis (2003) found a similar pattern: although grade retention in any grade was found to be related to lower enrolments in higher education, students retained in grade 6 through 8 had the lowest odds of enrolment.

However, the studies of Pustjens et al. (2004) and Ou and Reynolds (2010) have some shortcomings. Pustjens et al. (2004) concluded that grade retention was negatively related to both enrolment and success in higher education, but this conclusion can be criticised because of the research design of the study, and especially the comparability of grade retainees and continuously promoted students (who seem to differ not only in terms of e.g. socioeconomic background previous achievement, but also in terms of interest in learning tasks, effort, attitudes and several other background variables (see further)). This comparability is a recurrent critique referring to a lot of previous studies on grade retention. Recently, more advanced techniques such as propensity score matching/stratification were used, leading to more credible conclusions (e.g. Ehmke et al., 2010; Hong & Raudenbush, 2005; Hong & Yu, 2008; Wu et al., 2008, 2010). To our knowledge, only one study – the study of Ou and Reynolds (2010) – focussed on the effects of grade retention on enrolment in higher education and used a more advanced statistical analysis. However, the major shortcoming of the study of Ou and Reynolds (2010) is that they merely focussed on enrolment and did not consider success in higher education.

In the present study, we combine the strengths of the studies of Ou and Reynolds (2010) and Pustjens et al. (2004) and will focus on the effect of grade retention on enrolment and success in higher education by making use of a propensity score matching.

RESEARCH QUESTIONS

In line with previous research on the relation between grade retention and higher education, we will extend the focus of the studies of Fine and Davis (2003), Ou and

Reynolds (2010) and Pustjens et al. (2004) on enrolment in higher education as dependent variables, by our focus on success in higher education. Following research questions are at stake:

1. Is there a difference between retainees and non-retainees concerning the enrolment in higher education?
2. Is there a difference between retainees and non-retainees concerning the enrolment in the different types of (higher) education?
3. Is there a difference between retainees and non-retainees concerning the results at the end of the first year in higher education (short term success)?
4. Is there a difference between retainees and non-retainees concerning the results at the end of the third year in higher education (long term success)?

METHOD

Data

For this study, we made use of data stemming from the Flemish longitudinal 'LOSO' project (LOSO = Longitudinal Research Project in Secondary Education) (Van Damme et al., 2002). In this longitudinal project, a cohort of 6,411 students in 57 secondary schools was followed through secondary education and beyond. Students were tested on several aspects of achievement and non-cognitive outcomes on a regular basis. Additionally, the greater part of these students were followed during their course in higher education or their first years on the labour market. In this way, the LOSO project kept track of about 90% of the students of the original cohort, for 9 to 10 years in total. For the current study, we selected a subsample of students which met two criteria. First, we only selected students who graduated in the 6th year of the general track, the technical track or the artistic track (students of the vocational track were excluded). As can be seen in Figure 13, a student is allowed to enrol in all forms of higher education when he/she finishes secondary education in the general, technical or artistic track after the 6th year of secondary education. However, when a student finishes the 6th year of secondary education in the vocational track, he will not receive a diploma but a certificate of secondary

education, which does not give access to any kind of higher education. In order to participate in higher education, these vocational graduates need to follow an additional year in secondary education, giving them the opportunity to obtain a

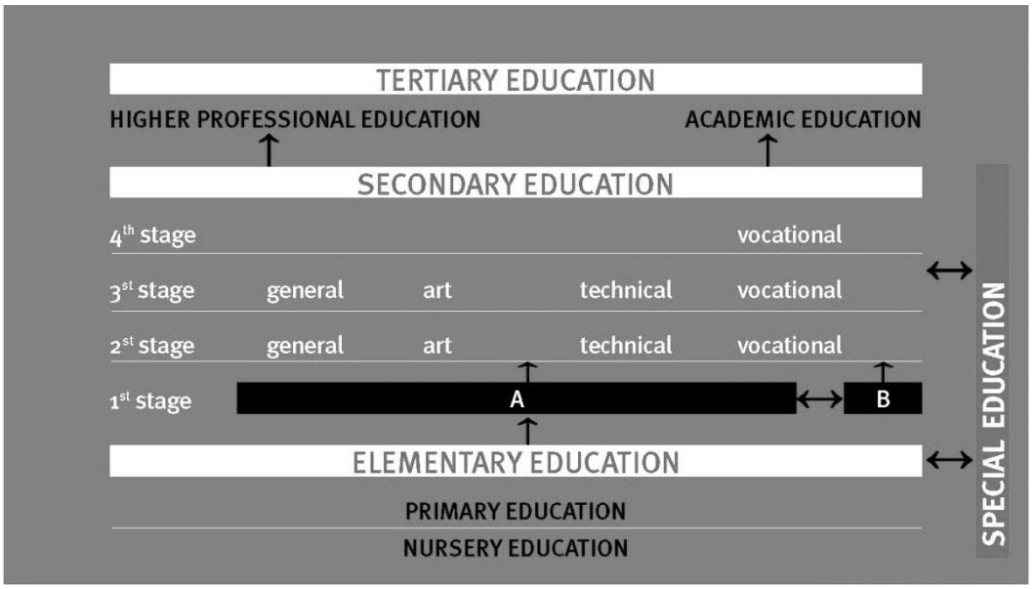


Figure 13: Secondary education in Flanders

diploma of secondary education, the required 'entrance ticket' to higher education. Excluding the group of students graduating in the 6th year of the vocational track from further analyses seems justified because it should be clear that the vocational track prepares for the labour market, rather than for higher education. In addition to this, only a minority of students graduating in the 6th year of vocational studies completes a 7th year and participates in higher education (in LOSO: only 6.7%). Next to the exclusion of vocational graduates, we also excluded early school leavers from further analysis.

This resulted in a final sample of 4,003 students, of which 3,166 enrolled in higher education. This 49.4% of the original LOSO-cohort enrolled in higher education is to a certain extent comparable to the ratio of students enrolled in higher education in the whole Flemish region (in 1999[xi]: 46.6% (EUROSTAT, 2013)).

Grade retention

Since the primary goal of the LOSO-project was to register school trajectories, we were able to rigorously determine in which grade in secondary school students were retained (in Flanders, this can also be the *choice* of a student/parent). Based on the grade, we made a distinction between early grade retention and late grade retention. Unlike Ou and Reynolds, who referred to early grade retention as grade retention in primary education, we defined early grade retention as retained in Grade 7-9 (for which we used in our analyses a dummy indicator, called RET_E). In addition, late grade retention refers to students retained in Grade 10-12 (RET_L).

Outcome

In this study, we consider three main outcome variables related to the pathway of the students after leaving secondary education: enrolment, short-term success and long-term success.

The first outcome considers the general enrolment in higher education. Next to the question whether or not a student enrols in higher education, we will further look at the type of higher education in which he or she starts. In general, a student can choose between three types of higher education: professional higher education college (one cycle, resulting in a professional bachelor degree; HEC_PROF), academic higher education college (two cycles, resulting in academic bachelor degree and master degree; HEC_ACAD), university (two cycles, resulting in academic bachelor degree and master degree; UNIV). Apart from higher education, a student can also choose to enrol in other forms of education (e.g. specialisation years in secondary education or on-the-job training; SPEC) or to start working (WORK).

Based on the LOSO data, we were able to determine if a student was successful in the short term (after 1 year) and in the long term (after 3 years). Success in the short term was defined as passing all the examinations in the first year (with or without a resit, if such should be the case). Success in the long term was defined as passing the three first years without delay. After 3 successful years, the student can be considered as a 'Bachelor' on the academic level or on a professional level, although at the time of data collection, the Bachelor-Master

structure was not yet introduced in Flanders (it was only introduced in 2004-2005). However, for reasons of comparability with other research on higher education, we opted for the term 'Bachelor's degree' when we talk about the end of the third year of higher education.

Matching and propensity scores

Research on the effects of grade retention usually suffers from the problem of possible selection bias. If we want to evaluate the effect of a certain treatment properly (in this case: grade retention), we would need to have two outcomes per subject: one where a student received the treatment and one where the same student did not receive the treatment. Of course this is not possible because the counterfactual of a specific outcome is not observed. An intuitive solution for this is to compare the average outcome of the treatment group with the average outcome of the non-treatment group. However, in doing this, one ignores initial differences between the two groups that existed even before the treatment. In our case, previous research indicated that several variables are related to grade retention. Apart from (prior) achievement variables (and variables related to achievement such as gender and socioeconomic status), several studies also reported relations between grade retention and parental involvement, well-being of the student and interest of the student in the subject matter. To properly investigate the effect of our treatment, we would like to have a treatment group which is comparable to the control group on all relevant pre-treatment covariates. If we would match individuals who received the treatment to individuals who did not receive the treatment, and use these matched individuals in further analysis, we can avoid selection bias since individuals in both groups are similar in all relevant pre-treatment characteristics (Caliendo & Kopeinig, 2008). This seems feasible when the number of pre-treatment covariates is rather limited, but once the number of pre-treatment covariates increases, matching on all these covariates becomes impossible. This is where a balancing score, such as a propensity score (PS), can be a solution. A PS is defined as the "…conditional probability of assignment to a particular treatment given a vector of observed covariates" (Rosenbaum & Rubin, 1983, p. 41). For the construction of the propensity score, we followed a stepwise procedure where we first selected the relevant covariates and used these to calculate the PS. In a next

step, this PS was used to match retained students to promoted students which had a comparable PS, and thus a comparable chance to be retained. After matching, the matched cases were used in multilevel regression analyses with the outcomes of interest.

For the first step, the selection of covariates, we followed the guidelines proposed by Imbens and Rubin (2012). As such, only predictors that were theoretically related to the treatment and the outcome were included. Rubin (2008) emphasizes the importance of the theoretical relation rather than only an empirical relation in the data, specifically for the relation between a covariate and the outcome. The idea behind PS-matching is to mimic a randomized controlled trial (RTC) as closely as possible. In an RTC the researcher has no idea (and access) to the outcome of the RTC. Therefore, Pattanayak, Rubin and Zell (2011) even suggest to *hide* the outcome data, "…mirroring the separation of study design and outcome analysis in randomized experiments" (p. 897). Based on theoretical grounds, we selected for early grade retention a set of 138 variables. For late grade retention, we selected 162 variables[xii]. These variables were all related to grade retention and/or achievement in general. For late grade retention we had more variables, because we could also correct for achievement in the early years of secondary education.

Although Imbens and Rubin (2012) agree with the theoretical relation between pre-treatment covariates and the treatment/outcome, they advise the use of a stepwise and iterative procedure to select the most *relevant* pre-treatment variables in relation to the treatment. This is done by adding covariates in a logistic regression model (with the treatment as a dependent variable) one at the time and by evaluating the model fit. If the addition of a covariate improved the model fit, this covariate was kept in the subsequent regression model wherein we added another covariate and evaluated the model fit again. This procedure was repeated until the addition of a new covariate did not improve the model fit anymore (once a covariate was included in the model, it stayed in the model even if in a later stage its effect was not significant anymore).

In the second step, we used the selected pre-treatment variables to compute the PS. This PS was then used in the matching procedure, where we matched students of the treatment group to students with a similar PS in the non-treatment group. In previous grade retention research that used PS scores, the matching was done by either stratification (Hong & Yu, 2008) or *k:*1 matching (Moser et al., 2012; Wu et al., 2010). When one uses a stratification approach, students with a similar PS in the treatment and control group are grouped together in subgroups. These subgroups are constructed by taking, for example, the quintiles of the PS distribution. Although this approach seems tempting and easy to use, it has a serious drawback: it is not clear how many subgroups are needed to remove the initial bias. Rosenbaum and Rubin (1983) stated that 5 subgroups are enough to remove 90% of the initial bias, but Lunceford and Davidian (2003) found that – especially in larger samples – 5 subgroups were insufficient to remove most bias. Up till now, there are few guidelines for the optimal number of subgroups, making it less attractive or reliable to use subgroups in combination with PS. The other popular matching approach, *k*:1 matching ('nearest neighbour matching'), also has some drawbacks. This *k*:1 matching can lead to very poor matches, because many comparison units are discarded. In its most simplest form – 1:1 matching (matching of 1 treated with 1 control unit) – this can lead to a large reduction of the sample size. Although this is not always problematic, depending on the number of treated units that are discarded (see Ho et al., 2007), there are more appropriate and sophisticated matching algorithms available.

As an alternative to these two popular approaches, Rosenbaum (1991) introduces the *full matching*, which keeps the middle between *k*:1 matching and subclassification (subgroups). With full matching, the number of subclasses is automatically determined and every subclass contains at least 1 grade retainee and 1 promoted student, with minimal distance between their respective PS. In every subclass, students received a weight: retained students received a weight of '1' and promoted students received a weight equal to the number of retained students in the subgroup divided by the number of promoted students (for more information, see Stuart and Green, 2008). We evaluated the quality of the matching by looking at the standardized bias, computed as

$$\frac{\bar{X}_{ret} - \bar{X}_{prom}}{\sigma_{ret}}$$

with $\bar{X}_{ret} - \bar{X}_{prom}$ representing the difference in means of each covariate between the retained and promoted students, and σ_{ret} the standard deviation in the original full comparison group. Standardized biases smaller than |.25| indicate a good match and balance in covariates between the two groups (Ho et al., 2007). If this balance was not achieved, higher order terms or interactions were included (or the variable was excluded from further analysis).

In this study we calculated four different propensity scores: the first PS is used to match early retained students with promoted students and uses the whole sample. We use this first PS to predict higher education entry. The second PS is used for evaluating the effect of early grade retention on success in higher education and therefore only focuses on the sample of students entering higher education. The third and fourth PS follow the same logic, but for late retained students.

Analysis

In the analysis, we used the weights produced by the full matching procedure in the previous step, and included these weights into a regression model. Because students are nested within schools, we also take into account the multilevel structure of the data. The dependent variable in the first, third and fourth research question is binary, so we applied a weighted multilevel logistic model.

For the second research question, we used a weighted multinomial logistic regression with the academic track (UNIV) as the reference category. We used a single-level multinomial logistic regression (ignoring the school effects) instead of a two-level model, because of computational issues of multinomial logistic regression in combination with the use of weights[xiii].

RESULTS

In the following paragraphs, we present the results of the matching procedure and the effect of early and late grade retention on the dependent variables. We will always make a clear distinction between early and late retention, and will report the results likewise.

Before turning to the main results, we first discuss the quality of the different matching's. As mentioned in the previous section, we calculated 4 different propensity scores and used these to match the different groups. Table 14 gives an overview of the number of matched students, and the number of covariates used in the different matching's.

Table 14: number of matched students and variables used

	Control	Treated (retained)	Number of covariates in matching
Early retention, total group			$n_{early,total}$=23
All	3,529	473	
Matched	3,343	447	
Discarded [a]	186	26	
Early retention, participation in higher education			$n_{early,higher}$=18
All	2,907	259	
Matched	2,503	242	
Discarded [a]	404	17	
Late retention, total group			$n_{late,total}$=31
All	2,622	548	
Matched	2,404	544	
Discarded [a]	218	4	
Late retention, participation in higher education			$n_{late,higher}$=28
All	2,285	371	
Matched	2,121	370	
Discarded [a]	164	1	

[a] *students with a PS outside the region of common support (i.e. PS that had no overlap in the other group) were discarded from further analysis*

The PS of early retention for the whole sample, was based on 24 covariates, and the PS for the sample enrolled in higher education was based on 18 covariates. The PS of late retention for the whole was calculated by taking into account 31 covariates, and for the group of students enrolled in higher education, we took 28 covariates into account. For further details of the matching, we refer to Appendix 3. In general, all the matching's resulted in standardised biases lower than |.25|, indicating a good balance.

We present the effect of early retention on the different outcomes mentioned in our research questions, followed by the results of the effect of late retention. These effects are easy to interpret: as we already matched on several background variables, we only needed to report the variable referring to the 'treatment', in our case early/late grade retention. As such, the reported value can be considered as the 'real' effect of grade retention.

Early retention

First of all, we were interested in the participation to higher education; do students retained in grade participate less in higher education? And if they participate, do they make less ambitious choices regarding the type of institution they enrol into? As the results in Table 15 indicate, the participation rate of early grade retainees in higher education is significantly lower $\left(\beta_{1,part\_early} = -0.594, p < .001; OR = .55,\right)$ compared to students who proceeded normally through secondary school. Students retained in Grade 7-9 have 45% lower odds to participate in higher education, compared to promoted students.

Table 15: effects of early retention on enrolment and success

			Est.	SE/SD[a]
Participation				
	Fixed			
	Intercept	$\beta_{0,part\_early}$	3.699***	0.939
	RET_E	$\beta_{1,part\_early}$	-0.594***	0.163
	Random			
	Schools		1.737	1.318
Short-term success				
	Fixed			
	Intercept	$\beta_{0,short\_early}$	-0.858	0.645
	RET_E	$\beta_{1,short\_early}$	-0.138	0.163
	Random			
	Schools		1.186	1.089
Long-term success				
	Fixed			
	Intercept	$\beta_{0,long\_early}$	-0.966	0.710
	RET_E	$\beta_{1,long\_early}$	-0.009	0.173
	Random			
	Schools		1.562	1.250

*Note: * $p < .05$; ** $p < .01$; *** $p < .001$*
[a] For technical reasons, the lme4-package (Bates, Maechler, & Bolker, 2013) in R gives standard deviations for variance estimates, instead of – the more common – standard errors

When we investigate the decisions made after secondary education more in detail (see Table 16), we find that early grade retainees choose more for a direct entry into the labour market ($\beta_{RET\_E,WORK} = 0.789, p = .005, OR = 2.2$); the odds of going directly to the labour market are 120% higher than for students who were never retained in grade. There are no significant differences in enrolment between the different types of higher education (professional higher education college, academic higher education college or university). The difference in enrolment

between grade retainees and promoted students in a specialisation year is not significant.

Table 16: effect of early retention on participation in higher education: multinomial analysis

		Intercepts			RET_E	
		Est.	SE		Est.	SE
Reference: UNIV						
HEC_PROF	β_{HEC_PROF}	2.965*	1.224	β_{RET_E,HEC_PROF}	0.099	0.259
HEC_ACAD	β_{HEC_ACAD}	-3.327*	1.669	β_{RET_E,HEC_ACAD}	0.084	0.331
SPEC	β_{SPEC}	2.464	1.665	$\beta_{RET\_E,SPEC}$	0.029	0.350
WORK	β_{WORK}	-2.103	1.403	$\beta_{RET\_E,WORK}$	0.789**	0.279

*Note: * p < .05; ** p < .01; *** p < .001*

In a next step, we were interested in the success of early grade retainees in higher education. To properly investigate the effect of grade retention, we calculated new propensity scores for those students entering higher education, which was a more restricted sample. With the new propensity scores and the corresponding weights, we were able to judge a student's success, both in the short-term as well as in the long-term. Based on the results in Table 15, there seems to be no significant effect of early grade retention on success in higher education, neither in the short-term, nor in the long-term.

Late grade retention

In presenting the results of late grade retention, we follow the same order as for the effects of early grade retention. First, when we investigate the difference in higher education participation between late retainees and continuously promoted students (Table 17), we see that students retained in the last three years of secondary education participate significantly less in higher education $(\beta_{1,part\_late} = -0.404, p = .007, OR = .67)$. The odds ratio of 0.67 indicates that students retained in Grade 10-12 have 33% lower odds to enrol in higher education, compared to promoted students.

Table 17: effects of late retention on enrolment and success

		Est.	SE/SD[a]
Participation			
Fixed			
Intercept	$\beta_{0,\text{part late}}$	-2.881	3.058
RET_L	$\beta_{1,\text{part late}}$	-0.404**	0.150
Random			
Schools		1.010	1.005
Short-term success			
Fixed			
Intercept	$\beta_{0,\text{short late}}$	1.591	2.144
RET_L	$\beta_{1,\text{short late}}$	-0.392**	0.129
Random			
Schools		0.746	0.864
Long-term success			
Fixed			
Intercept	$\beta_{0,long\ late}$	0.540	2.195
RET_L	$\beta_{1,long\ late}$	-0.526***	0.138
Random			
Schools		0.275	0.525

*Note: * $p < .05$; ** $p < .01$; *** $p < .001$*

[a] *For technical reasons, the lme4-package (Bates, Maechler, & Bolker, 2013) in R gives standard deviations for variance estimates, instead of – the more common – standard errors*

Based on a further (multinomial) analysis (Table 18), it seems that students retained in the last three years are significantly more inclined to opt for a direct entry into the labour market $\left(\beta_{RET\_L,WORK} = 1.663, p < .001, OR = 5.28\right)$, compared to enrolment in university education. The odds of going to the labour market instead of university education, are five times higher for late retainees compared to promoted students. They also enrol significantly more into a professional higher education college $\left(\beta_{RET\_L,HEC\_PROF} = 1.097, p < .001, OR = 3.0\right)$ or an academic higher education college $\left(\beta_{RET\_L,HEC\_ACAD} = 0.489, p < .001, OR = 1.63\right)$, compared to university education.

Table 18: effect of early grade retention on participation in higher education: multinomial analysis

		Intercepts			RET_L	
		Est.	SE		Est.	SE
Reference: UNIV						
HEC_PROF	β_{HEC_PROF}	4.858***	1.432	β_{RET_L,HEC_PROF}	1.097***	0.186
HEC_ACAD	β_{HEC_ACAD}	1.036	0.763	β_{RET_L,HEC_ACAD}	0.489*	0.228
SPEC	β_{SPEC}	9.991***	0.925	$\beta_{RET\_L,SPEC}$	1.067***	0.269
WORK	β_{WORK}	4.128***	1.067	$\beta_{RET\_L,WORK}$	1.663***	0.206

*Note: * $p < .05$; ** $p < .01$; *** $p < .001$*

Analogous to the analysis of success after early grade retention, we recalculated the propensity score for late grade retention for the students in the more restricted sample. Based on this new propensity score and corresponding weight, we examined the effect of late grade retention on short-term success and long-term success (Table 17). We found a negative effect of late grade retention for both success outcomes: compared to continuously promoted students, students retained in the last three years of secondary education had 32% less chance to pass all examinations and were significantly less successful in the short-term $(\beta_{1,short\_late} = -0.392, p = .002, OR = 0.68)$. The same conclusion holds true for success in the long-term: late retainees had a significantly lower chance to obtain their Bachelor's degree by the end of the third year $(\beta_{1,long\_late} = -0.526, p < .001, OR = 0.59)$.

DISCUSSION

In this study, we were interested in the relation between grade retention and enrolment/success in higher education. This focus on higher education was important for several reasons. First, participation in – and successful completion of – higher education is one of the main targets of the ambitious European EU2020 plan. Second, a diploma of tertiary education still leads to better outcomes in later life: improved chances for employment, a reduced risk of unemployment and higher earnings (OECD, 2013)

As suggested by previous studies, the odds for a student to enrol in higher education seem to decrease when this student repeated a grade throughout his educational career. In the previous sections, we discussed some of the shortcomings of these studies. Basically, these shortcomings were threefold: (1) Only effects of grade retention in primary education or in the first grades of secondary education were examined (Fine & Davis, 2003; Ou & Reynolds, 2010), (2) They did not match grade retainees with comparable but normally promoted students (Fine & Davis, 2003; Pustjens et al., 2004) and/or (3) They only considered enrolment and did not make any notice of success in higher education (Fine & Davis, 2003; Ou & Reynolds, 2010). In our study, we made use of a propensity score matching approach for both early grade retainees (Grade 7 – Grade 9) and late grade retainees (Grade 10 – Grade 12), and we considered the effect of both timings of retention on participation and short/long term success in higher education. Following this matching, we could compare students retained in grade with promoted students but with a similar chance to be retained.

As for the first research question, we can confirm the results of previous studies: the entrance into higher education seems to be a first sorting between grade retainees and promoted students. However, we cannot explain the difference in entrance rates in terms of the well-established link between grade retention and dropout, as Ou and Reynolds (partially) do. Because we only selected those students who actually graduated from secondary education, our effect is – in a way – a more 'pure' effect of grade retention on participation in higher education. Where the timing of grade retention seems to make no difference when it comes to mere entrance into higher education, we do find a difference in the choice of type of higher education depending on the timing of grade retention. Early retainees do not make less ambitious choices and enter universities to the same extent as promoted students. In contrast, when students were hold back in the last three years of secondary education, this seems to have implications for their further choice of study program in higher education, where late retainees are less inclined to opt for university education. This result brings some nuance to the conclusion of Pustjens et al. (2004), where the difference between early and late grade retention was not made. Based on our results, it seems that the 'unambitious choice' of grade

retainees in Pustjens et al., can be explained by the specific choices of the late retainees. In that sense, later grade retention seems to have a more severe impact compared to early grade retention.

The question remains though, why do retained students participate less in higher education? A sound answer to this question cannot be derived from our results, but based on previous research, we believe that a plausible explanation can be found in the noncognitive characteristics of students. As Ou and Reynolds (2009) and Heckman, Stixrud and Urzua (2006) state, these noncognitive characteristics sometimes have a stronger effect on college attendance than cognitive outcomes. In the context of grade retention, one of the more interesting characteristics is the engagement of the student, because students with lower levels of engagement are found to participate less in higher education (Horn & Carrol, 1997). In line with Finn's 'participation-identification' model (1989), grade retention can lead to lower levels of identification/engagement with school. Evidence for this process was found by Martin (2011) who concluded that grade retention was a significant negative predictor of homework completion (as an indicator of engagement). One can imagine that, when a student then eventually graduates, he has such a low level of engagement towards school, that participating in higher education is not even worth consideration.

When it comes to the effect of early/late retention on short/long term success, we also find differences related to the timing of retention. Especially late retention has a strong negative effect on short- and long-term success in higher education. Based on that result, one might conclude that early retention is *less* harmful for students. This, however, would be an overly simplistic conclusion given the lower chance to enrol in higher education. The lower success rate of late retainees was somewhat surprising, especially because these retainees made less ambitious choices, starting more in the less prestigious (simplistic: less difficult) tracks. For an interpretation of this effect, one should keep in mind that these late retainees were also those students struggling in the last years of secondary education, just before entering higher education, and this struggle seems to persist up until higher education. Moreover, students retained in the last years of secondary education struggled with content which was more comparable to that of higher

education. This may explain some of the differences in success between early and late retention. Furthermore, if we look at (the scarce) previous research on the effects of grade retention in secondary education on the achievement in the subsequent years, there is some additional evidence for this 'timing' effect of retention. For students retained in Grade 9 – which corresponds to 'early retention' in our sample – there seems to be no effect of grade retention on future achievement at all (Ehmke et al., 2009). Recently, another study (Lamote, Pinxten, Van Den Noortgate & Van Damme, 2013) focussed on the effect of grade retention in Grade 8 and found the opposite effect of Ehmke et al. (2009): holding a student back in Grade 8 resulted in a decreasing achievement, from which they concluded that also early grade retention was harmful for students. It is important to note that this result does not contradict our results: in Lamote, Pinxten et al. (2013), only 14% of the grade retainees graduated in the academic track of secondary education while a lot of students graduated in the vocational track. In the present study, we deleted students that finished secondary education in the vocational track, resulting in a more restricted sample (but more relevant for our analyses), making it less comparable to our earlier results. In contrast to Ehmke et al. (2009) and Lamote, Pinxten et al. (2013), Uysal (2010) focussed on the effect of retention in Grade 10 – corresponding to late retention in the current study – and concluded that holding a student back at such a late point in their educational career, has a strong negative effect on further achievement.

But one may wonder, is this all an effect of just 'timing', or is it more the rationale behind the decision to repeat a grade? Can it be that students retained in the first years of secondary education did not master basic content or were not in the right track, and were therefore advised to repeat a year in order to 'catch up', while students retained in the last years of secondary education did not master the content relevant for higher education, with consequences on their success in higher education? Is it that – as also noted by Fruehwirth, Navarro and Takahashi (2011) – simply different types of students are retained at different grades, because of different thresholds in every grade? Although we control and match for several background characteristics, we had no information on the 'reason' the teacher gave for the decision to retain or the different thresholds across the grades. And if we

would have that information, we would need to incorporate it into a dynamic selection model where we, ideally, estimate the effect of grade retention for every grade.

This lack of information brings us to the first limitation of our study: we do not have enough information in the LOSO-database to correct for grade-specific student characteristics, necessary to calculate a correct propensity score, needed for matching students. A second limitation is the availability of data on the success of students. As mentioned in the outcome section, we only have reliable data on the success/failure of a student, 3 years after he/she enrolled in higher education. Hence, we cannot say anything about the period afterwards, which may be important for unsuccessful students. It may be the case that these students do obtain a Bachelor's degree, but after one (or more) years of delay. Therefore, in our study, the 'unsuccessfulness' in the long term should be interpreted with some caution.

CONCLUSION

From our results, we can conclude that grade retention in secondary education affects not only the chances to enrol in higher education, but also the type of higher education chosen. When a retained student does enrol in higher education, it depends on the moment when he was retained, whether he will be successful or not. So, can we really call these students 'successful failures' as Ferguson et al. (2001) (partially) did, just because they graduated from secondary education? In our opinion, no, because of the significantly lower chance in this group for enrolling in higher education and (for late retainees) the lower success rates in higher education. On the other hand, do we need to abandon the practice of retention, because of these implications? Not necessarily, because our results are only mean effects, and some students do well by an extra year in the same grade. A well-thought decision for every individual student, taking into account all the pros and cons of grade retention seems to be the best advice

Chapter 6

General discussion

More than ever, a diploma of secondary education is (one of) the preferred entrance ticket(s) for the labour market. Still, more than one out of ten students does not reach this level and leaves school prematurely. We label these students as 'early school leavers', 'dropouts' or 'unqualified school leavers'. Throughout this dissertation, we analysed the relation between several predictors and early school leaving. In this section, we summarize the main findings of the preceding chapters, followed by an overview of the main limitations of our studies, which act as a starting point for suggestions for future research. We end this dissertation with implications for policy and practice.

EARLY SCHOOL LEAVING: METHODOLOGICAL CONSIDERATIONS AND THE EFFECT OF ENGAGEMENT

In the first chapters of our dissertation, we were interested in the idea of 'early school leaving' in general, and more specifically in two important shortcomings of previous research. First, previous research did not always account for the longitudinal nature of dropout-data (and the problems related to this longitudinal character), but modelled dropout at the end of a certain period (e.g. at the end of compulsory education (Archambault et al., 2009; Janosz et al., 2008)). In this dissertation, we were interested not only in the occurrence of an event, but also in the timing of the event. Therefore, we followed the recommendations of Singer and Willett (2003) and conducted a discrete-time survival analysis (DTS). In Chapter 2, we explored different approaches in working with this DTS, and in that sense, Chapter 2 can be considered as an illustration of different – common and less common – methodological approaches of analysing early school leaving, and longitudinal (multilevel) data in general. As we were interested in the effect of school level characteristics on early school leaving, we opted for a multilevel approach, where the variance can be decomposed and assigned to the different levels. In an educational setting, this multilevel approach looks straightforward since students are nested within schools, so variance can be allocated to the student level and the school level. Things get more complicated when a researcher is interested in longitudinal analyses, because it is not uncommon that students change

schools. There are different approaches to handle this student mobility: a researcher can simply ignore these school changers and take the first/last school as a basis for clustering, or, even worse, can simply delete students that changed school. However, the appropriate way to handle these mobile students, is by working with a cross-classified or multiple membership approach (Fielding & Goldstein, 2006; Goldstein, 2011, Raudenbush & Bryck, 2003). In this chapter, we compared the effect of different approaches on the parameter estimates, the standard errors of the parameter estimates, and the between school variance. The different models we tested were (1) a model without considering the school level, (2) a model accounting for the clustering of students based on the first school, (3) a model accounting for the clustering based on the last school, (4) a multiple membership model, (5) a cross-classified model, and (6) a model where we ignored all the students that changed schools. Although the model without considering the school level is often used in research on early school leaving (Bowers, 2010; Gesthuizen et al., 2005; Roderick, 1994), we conclude that this model is invalid since the use of multilevel models (no matter which kind of clustering is used) indicate that there are differences between schools which should be taken into account by assigning variance to the school level. Next, we compared the second, third and sixth modelling approach, because these are the most common (and pragmatic) ways of handling student mobility (by frankly ignoring the fact that students change schools). Of these three approaches, the clustering based on the last school seems to yield the best model fit (a reader might conclude that the low DIC of the 'delete approach' indicates a better fit, but this model is impossible to compare with the other models because the sample of this model differed (too much) from the sample used in the previous models). The fact that the multiple membership approach yielded a higher DIC than the last approach, came as a surprise to us. Probably, this is due to the weighting scheme we used, where we assumed that every school attended has the same impact, so we calculated the weights as the proportional time spent in the school, irrespective from the order of attendance. However, Fielding and Goldstein (2006) suggest that the more recent schools have a higher impact, and weights should be calculated in accordance with this. We could test different weighting schemes, but this brings along practical difficulties (very time consuming to find a 'proper' scheme). Instead of experimenting with different weighting

schemes, we turned over to a cross-classified approach, which yielded the best model fit. In a next step, we extended the 'best pragmatic model' (last school) and the cross-classified model with student and school characteristics, and concluded that ignoring this cross-classification can have severe consequences on the standard errors and parameter estimates, which can (potentially) lead to incorrect conclusions concerning the effect of student or school characteristics on dropout. From this results, we would advise researchers working with longitudinal data to handle student mobility in the appropriate way, or, if not feasible, to use a good pragmatic solution always keeping in mind the possible effects on the estimates and consequently, on the conclusions. Apart from the methodological conclusions, this study can also act as a first framework for early school leaving in Flanders and confirmed some of the international tendencies. With our model, we can conclude that the highest dropout rate is in Grade 11, and that several student and school characteristics significantly predict the likelihood for early school leaving. Although interesting in itself, the predictors of early school leaving in this model are difficult to deal with, especially by an individual teacher in a classroom setting. Therefore, we focussed in a next chapter on a more malleable construct in tackling early school leaving: school engagement. Chapter 3 was devoted to the effect of the development of engagement on early school leaving. Although this was also the topic of some previous studies (Archambault et al., 2009; Janosz et al., 2008), we extended these studies by our explicit focus on the moment of dropout. By using a discrete-time survival mixture (DTSM) model, we were able to model different trajectories of engagement (mixture part) and to relate these trajectories to the risk of dropping out in every grade (survival part). The latter is the main difference between our study and the previous studies. Ideally, we would account for the student mobility by making use of the cross-classified approach that we explored in the previous chapter, in combination with a DTSM. Unfortunately, combining cross-classifications with a mixture model in a latent framework seems to be impossible with the software we used (Chen, 2012) which forced us to use the 'last school approach'.

We conducted our DTSM on two indicators of engagement: emotional engagement (measured by student-teacher relations) and behavioural engagement

(measured by the attitudes towards homework). For both engagement constructs, more than half of the students started at a high level of engagement and followed a (relatively) stable pattern. Students in this trajectory encountered few problems concerning graduation: most of these students obtained a diploma of secondary education (or certificate of the 6th year of vocational education), sometimes with one (or more) years of delay. So, it is not this group that should be the subject of concern. For behavioural as well as for emotional engagement, there are two more 'problematic' groups: students in trajectories that follow a steep decline or students that already start at a very low level. It seems that students in these declining or low trajectories have a higher chance of dropout, and even more, the probability of dropout already increases starting from Grade 9.

Looking at these trajectories of engagement, one may wonder: is there no group of students that follows an upwards development, an increase in engagement over time? No, but this came not as a surprise, given the findings of several previous studies on engagement. Van de gaer et al. (2009) reported for both emotional and behavioural engagement, on average, a decline over time, and, as we mentioned in Chapter 3, if we look into the indicators of both engagement constructs, the decrease over time also makes perfect sense (see e.g. Opdenakker et al., 2012; Xu, 2004). However, these authors (Opdenakker et al., 2012, Van de gaer et al., 2009; Xu, 2004) did not use a mixture approach, so they were only able to capture the average trajectory of engagement, which may cover a (small) group of students deviating from this declining trajectory. Only few studies used a mixture approach, and did found a group with an increasing engagement, but this group usually consisted of hardly 10% of the students. In the two previous studies considering the development of engagement (Archambault et al., 2009; Janosz et al., 2008), the researchers found two groups of students following an increasing pattern of engagement: a group which showed a transitory increase in the early years, and a group which showed a more stable increase. However, an increasing engagement did not protect the students from dropping out either. Janosz et al. (2008) concluded that, of both increasing engagement groups, students showing a strong but transitory increase in the early years had the highest chance to drop out. What is more, this group had even more dropouts than the group of students who

followed a decreasing engagement trajectory. It seems that students in this strong increasing group first 'fly high', but 'fall hard' and that, next to the level of engagement, also the stability of the level of engagement during the trajectory matters for avoiding dropout.

It is also interesting to look at the background characteristics of the students in the different trajectories, and especially the predictive power of gender on group membership. In Chapter 2, we confirmed the often cited relation between gender and early school leaving, but when we included engagement in Chapter 3, the direct effect of gender on dropout disappeared. Boys tend to have a higher chance to be member of the 'problematic' engagement groups (i.e. low, low & decreasing, or high & decreasing), and therefore, boys have a higher probability to dropout. So, does the effect of gender on dropout runs by engagement, and does engagement acts as a mediator between gender and dropout? This is plausible, but we cannot confirm this with our model. A second background characteristic of importance, is related to the pace at which a student proceeds through education. Students who repeated a grade once or more than once in primary or secondary education, had a higher probability to be in the 'problematic' engagement groups, and consequently, a higher chance to dropout. Especially for students retained in secondary education, the relation between engagement and grade retention seems to be a 'chicken-and-egg' problem. Indeed, the relation between these two elements can work in both directions: grade retention can cause a drop in engagement but a decrease in engagement can also be (one of) the cause(s) of grade retention. Again, this is a question that cannot be answered with our model.

GRADE RETENTION: EFFECTS IN THE SHORT-TERM AND IN THE LONG-TERM

From the previous chapters, it became clear that grade retention is related to early school leaving which suggest that grade retention might not be as harmless as a lot of teachers and parents think. Therefore, Chapter 4 and Chapter 5 are uniquely devoted to the effects of grade retention, where we considered different outcomes. From a methodological viewpoint, both studies shared one common problem: possible selection bias. To deal with this problem, we used a propensity score (PS) matching approach (Rosenbaum & Rubin, 1983). By using a PS matching approach, we try to mimic a randomized controlled trial (RCT) as much as possible. Such RCT's have a very strong internal validity, but setting up an RCT for the research questions on grade retention is not feasible for ethical and practical, which is why we use a PS-matching. Therefore, research working with propensity scores is often referred to as quasi-experimental research. The PS represents the probability to receive a certain treatment, given a large set of background variables. If one uses this PS to match students in the control condition with students in the treatment condition, the distribution of covariates defining the PS is the same in both groups (although it is possible that different values of the covariates can lead to a same PS). The idea behind a matching based on a PS, is that the effect of the treatment can be interpreted as a 'pure' effect, without selection bias. So, if, for example, prior achievement scores are included in the PS, two matched students are assumed to have the same distribution of achievement scores after the matching.

When using a PS matching approach in our own studies on grade retention, we considered different treatments: in Chapter 4 the treatment was 'retained in Grade 8', while in Chapter 5 we had two treatments: 'grade retention in Grade 7-9' and 'grade retention in Grade 10-12'. We selected several covariates related to these treatments (separately), computed the propensity to be retained and matched students in both conditions based on their propensity to be retained. The matched students were then used in subsequent analyses. These analyses covered outcomes on two educational levels: the secondary school level and the higher education level. Both levels are important in research on the effects of grade retention, because of the scarcity of studies covering these levels. Most of the time, grade

retention studies focused on the primary school level (e.g. Hong & Raudenbush, 2005; Moser, West, & Hughes, 2012; Wu et al., 2008). The results of these studies point out that grade retention is not always the best solution. Compared to students with an equal risk to be retained, but who were promoted instead, retained students only perform better in the short-term but this head start fully dissipated in the longer term (Moser et al., 2012; Wu et al., 2008). Regarding psychosocial functioning, conclusions of previous research in primary school often supported Festinger's social comparison theory (Festinger, 1954), namely that retained students feel more competent in the year of retention. This effect was, again, only observed in the short-term and disappeared the years after retention (Bonvin et al., 2008).

The question in our studies was: do these effects hold true in secondary education, and are there any differences between short- and long-term effects? Our first study on grade retention focused on the effect of being retained in a specific grade, on achievement and academic self-concept. Because of data restrictions, we could only analyse both short- and long-term effects of retention in Grade 8. Our results confirm most of the findings of grade retention in primary education. Overall, grade retention seems to have no (negative or positive) effect on achievement in the year of retention. When we split the treatment into conditions based on the certificate a student received beforehand, we see that students who received a B-certificate (i.e. a student can choose between repeating the grade or change over to another (lower) track) and changed over (so were promoted) already performed lower than students who received an A-certificate. Generally speaking, our results on the short-term effects of grade retention on achievement corroborates the findings of previous studies, indicating that these retained students are able to 'catch up' in the year of retention. This finding seems intuitive: because most of these grade retainees receive for the second time the subject matter they did not master, they will perform better because most of the subject matter is just repetition of the previous year. However, it is important to note that our conclusions are based on same-grade comparisons, where we compare the academic skills of the grade retainees in the year of retention (e.g. the second time in Grade 8, in our study in the school year 1992-1993) – with the cost of an extra year of schooling – with the skills of their younger, former classmates with a comparably poor performance a

year earlier (in Grade 8 in school year 1991-1992; see also Chapter 4). Most of the time, such kind of comparison is in favour of the grade retainees (Allen et al., 2009). Goos, Van Damme, Onghena, Petry and de Bilde (2013) found that, based on same-grade comparisons, grade retainees even outperform their promoted grade-mates. So, in our study, can we even call the performance of grade retainees 'catch-up' if they achieve at the same level the second time in Grade 8, as their promoted grade-mates (again, with a comparable poor performance a year earlier)? If this same-grade comparison favours grade retainees, we can expect that these retainees would score higher during the second year in Grade 8 and that way, take a head-start in order to deal with the subject matter the years after. Unfortunately, this is not the case in our study, making it difficult to state that grade retention has positive short-term effects.

In accordance with previous studies (Goos et al., 2013; Hong & Raudenbush, 2005; Hong & Yu, 2008; Moser et al., 2012; Wu et al., 2008), the effect of grade retention on achievement turns to be negative in the long-term. In our case, it is also interesting to look at the growth curves of the students receiving a B- or C-certificate. It seems that following the advice of the teachers to change track, instead of repeating a grade in the same track, is the best decision. Ironically, the poor performance of the grade retainees can also be (partially) explained by changing tracks. At the end of secondary education, the group of grade retainees has only 14% of students in the academic track and the majority is in the vocational track, where language education has a less important place in the curriculum, which can partially explain the large drop in achievement in Grade 12 (we only looked at the hours of language education, and did not consider other course subjects such as math).

Next to achievement, we also analysed the effect of grade retention on academic self-concept. Regarding the effect of grade retention in the short-term, our results are to some extent in support of the social comparison theory (and the Big-fish-little-pond effect): in the retention year, grade retainees score higher on academic self-concept compared to their former classmates when they were in Grade 8, most probably because they receive the subject matter for the second time and hence, feel more competent. It may also be the case that they compare their

performance (or the effort needed to perform) with their class-mates and perceive themselves as being more competent (Big-fish-little-pond). In the long-term, we can conclude that academic self-concept remains rather stable (although retained students score slightly higher by the end of Grade 12). When we look at the certificate the student received at the end of (the first time in) Grade 8, and focus on the long-term effects of this certificate, we can find some support for the conclusions of Wouters et al. (2012) that changing to another (lower) track boosts a student's academic self-concept.

As we discussed in Chapter 5, grade retention not only affects outcomes on the secondary education level, but has also implications on post-secondary outcomes. To analyse the effects of grade retention in secondary education on post-secondary outcomes, we followed an approach, similar to the approach in Chapter 4. However, in order to draw conclusions on the timing of retention, we made two groups of students: students retained in Grade 7-9 (early retainees) and students retained in Grade 10-12 (late retainees). We considered three outcomes, all related to higher education: enrolment in higher education, success in the short-term and success in the long-term. For both groups of retainees, we calculated two propensity scores for grade retention: one for the whole sample at the end of secondary education, and one for those students who enrolled in higher education. Doing this allowed us to draw better conclusions on the effect on success in higher education. If we would have only calculated a PS for the whole sample, and use this to draw conclusions on success in higher education, the effect of grade retention would be difficult to disentangle: is the effect on success a 'real' effect, and which part of the effect can be ascribed to the effect of grade retention on enrolment?

First, we focus on the effects of retention on enrolment in higher education. In line with previous studies (Fine & Davis, 2003; Ou & Reynolds, 2010; Pustjens et al., 2004), we found that both groups of retainees have a lower chance to enrol in higher education, but that only the group of late retainees makes less ambitious choices when they eventually enrol. Although we cannot fully confirm it with our data, we believe that the lower enrolment is due to a lower engagement of grade retainees. As we know from Chapter 3, students retained in grade are found in the more 'problematic' engagement groups (low or decreasing

groups), and following Finn and Zimmer (2012) this low engagement affects post-secondary outcomes such as enrolment in higher education. To put it bluntly: if a student was confronted with (repeated) setbacks in primary/secondary education, extending his educational career by enrolling in higher education is often not an option.

The differences between both groups of retainees become more pronounced once we consider the success of the students. It seems that repeating a grade at the end of secondary education has a more severe impact on success in higher education than repeating a grade at the beginning of secondary school. As we already discussed in Chapter 5, we do not see this difference as an effect of 'timing', but rather as an effect of the rationale behind the decision to hold a student back. Students retained in the last years of secondary education struggled with content that is more relevant for higher education, while students in the earlier grades struggled with more 'basic content' (for which they had more years to catch up). In our study, we corrected for pre-treatment characteristics (in line with the recommendations of PS-matching (Rosenbaum & Rubin, 1983)), so we did not include variables relating to the achievement in Grade 11 or Grade 12 (post-treatment). Ideally, we would estimate the effect of grade retention in every grade, but unfortunately, we do not have sufficient data to do so.

Now, what is the take-home message of the effects of grade retention on achievement, academic self-concept and post-secondary outcomes? Should we, given the negative effects of grade retention on achievement and enrolment in higher education, abandon the practice of grade retention? We would suggest not to completely abandon it, but always be aware of the possible implications and to consider every student's situation because one should keep in mind that our conclusions are based on average effects, and that some students may benefit from repeating a grade (e.g. students that were absent because of a long illness). On the other hand, a bulk of studies pointed at the negative effects of grade retention, indicating that the majority of retained students are worse off because of the grade retention.

STRENGTHS, LIMITATIONS AND SUGGESTIONS FOR FUTURE RESEARCH

Strengths

The studies presented in this dissertation have several strengths and limitations. The strength of this dissertation is in the use of state-of-the-art methodology in order to answer pressing, current educational questions. In almost every chapter, we turned away from the cross-sectional approach by using survival models and growth curve models. This yielded some new insights for educational theory and practice.

First, it became clear that using a cross-sectional approach to analyse early school leaving, covers the fact that students can drop out in several grades. Previous research (e.g. Archambault et al., 2009) often considered dropout as 'not in education or training at a certain age', but only focussed on the situation at that age and did not model the grade in which a student left school. Of course, one can say that students leaving school before the end of compulsory education are strictly seen as truants, and not dropouts so there is no need to consider them as such. But still, if one focuses on a certain age it is advisable to model the grade in which they dropped out, in order to reveal some tendencies. Because we modelled the grade of dropout in Chapter 2 and Chapter 3, we found that the dropout rate reaches the highest level in Grade 11. This may indicate that grade retainees leave education when they are legally permitted to, and do not stay in school one year longer in order to obtain a diploma. This phenomenon was also reported by Van Landeghem and Van Damme (2011).

Second, Chapter 3 was innovative in extending a survival model to a growth mixture model. Combining survival and mixture models not only resulted in different trajectories of engagement, but, more important, in different survival curves depending on the trajectory. By using this method, different aspects of the relation between engagement and dropout became clear. With this method, we can confirm Rumberger's statement (1987) that dropout is the final point of a process of disengagement. Moreover, we can conclude that students in lower and/or decreasing engagement groups have a higher probability to leave school without a

diploma, and, that the dropout of these groups starts very early in secondary education. This finding emphasizes that preventing early school leaving should start at the beginning of secondary school (and probably even in primary school), by identifying disengaged students.

Third, in evaluating the effect of grade retention, we considered both short-term as well as long-term effects by making use of a growth curve analysis. Considering effects in the long-term is proved to be important because, in Chapter 4, we corroborate the findings of previous studies (Alexander et al., 2003; Bonvin et al., 2008, Wu et al., 2008) that grade retention may be beneficial in the short-term, but turns out to be negative in the long-term, at least for students' achievements. The results in Chapter 5 underline the importance of the long-term aspect in grade retention research. As one of the few studies, we analysed post-secondary education outcomes of students retained in secondary education, from which it became clear that grade retention has an effect well beyond secondary education.

Apart from the longitudinal approach in the previous chapters, we accounted for (possible) selection bias in Chapter 4 and Chapter 5 by using a PS-matching. Using this PS-matching made it clear that a lot of covariates predict the likelihood to be retained, and should be taken into account.

Limitations and suggestions for future research

Throughout the different chapters, some limitations turned up and our conclusions should be interpreted with these limitations in mind.

A first limitation concerns the engagement scales we used, as these scales are more 'ad hoc' interpretations instead of validated scales. The engagement scales we used in Chapter 3 were based on the scales in the original survey (Smits & Vorst, 1982) and confirmed with a factor analysis on the 'well-being' questionnaire, which was administered in several grades. The scales of this survey were compared with the definition (and measurement recommendations) of the different dimensions of engagement, formulated by Fredricks et al. (2004). In addition to this, we also looked at the study of Van de gaer et al. (2009) and the scales she used to capture the different dimensions of engagement. This study was relevant because

Van de gaer also relied on the LOSO-database, and was faced with similar questions concerning the conceptualisation of engagement. Comparing the scales of the well-being questionnaire with the definition and conceptualisation of Fredricks et al. (2004) and Van de gaer et al. (2009) resulted in two indicators: one for behavioural engagement and one for emotional engagement. We found no indicator for cognitive engagement. Although we followed the suggestions of Fredricks et al. (2004) meticulously, the conceptualisation we used for both dimensions of engagement is debatable. Therefore, future research should make use of a validated engagement scale, in which the three dimensions of engagement are present. For an overview of possible instruments, we refer to Fredricks and McColskey (2012). Of course, we admit that assessing engagement is usually only a small part of surveys such as the LOSO-survey, and that researchers always have to find a good balance between the length of a survey and the content they want to cover. Nevertheless, because of the malleable nature of engagement, it is useful to include a validated engagement scale in future surveys, and what's more, there is no need to include a lengthy scale (for example: High School Survey of Student Engagement: 121 items) because shorter scales are also valid and have good reliability measures (for example: School Engagement Measure (Fredricks, Blumenfeld, Friedel, & Paris, 2005): 19 items).

A second limitation is related to the general character of some of the non-cognitive outcomes, in particular the academic self-concept scale and, again, the engagement scales. Both outcomes can be measured in two ways: as a broad outcome reflecting the feelings towards school in general, or subject specific, leaving more space for nuances depending on the subject. Concerning the academic self-concept, Marsh (2007) points at two (nearly) orthogonal factors: math self-concept (including self-concept of math, science and economics) and language self-concept (including self-concept of geography, history, native and foreign language). It may be surprising that both sub dimensions of self-concept are nearly uncorrelated, while achievement in math is usually highly correlated to language achievement. This can be explained by a combination of the well-known 'big-fish-little-pond' effect and the 'internal/external frame of reference' model, where a student compares his perceived abilities within the context of a reference group

(e.g. class-group) (see e.g. Parker, Marsh, Lüdtke, & Trautwein, 2013). With these relations in mind, it should be clear that a broad measure of academic self-concept is not very revealing, because it covers the differences caused by the subject under consideration. These domain-specific measures are also useful when one considers the engagement of a student. In contrast to the research on domain-specific self-concept, Eccles and Wang (2012) correctly note that the existing engagement measures are quite general and that domain-specific measures are really necessary. It is not difficult to imagine that a student feels more behaviourally engaged to one course subject than to another, which can have an effect on the achievement. In this context, Finn and Zimmer (2012) suggest to explore commonalities and differences between different subjects, and not just focus on one subject (e.g. math). Taken together, because of the broad, often insignificant, conclusion drawn from the general measures of self-concept and engagement, we would suggest future research to specifically focus on these domain-specific differences.

Our study is also limited in the available time-points. In the LOSO project, students who were not retained during their career in secondary education, have (in general) measures in four grades: Grade 7, 8, 10 and 12; students who retained a grade, have measurements one year later for Grade 8, 10 and 12. Although this is a very complete set of data, the lack of data on some time points had implications for the analyses in the different chapters. In Chapter 2 and Chapter 3, we used a discrete-time survival model, where we had a dropout measure for every grade. Because of the discrete-time approach, it is also very straightforward to include time-varying predictors, as this is one of the main advantages of a discrete time approach (Singer & Willett, 2003). Moreover, Bowers (2010) emphasizes that time-varying variables can predict dropout more accurate than time-invariant variables. Unfortunately, for most of the variables, it is not possible to include them as time-varying for every grade. For example, if we were able to include them, we would have included different achievement scores per grade, but in LOSO, we have no information on the achievement in Grade 9 and Grade 11. Therefore, we only included the achievement at the start of secondary education. The lack of information on certain time-points becomes more problematic in Chapter 4, because it places restrictions on our comparison groups. In Chapter 4, we compared the

scores of students retained in Grade 8 with the scores of students who were in the same grade, but one year earlier. This approach, also known as 'same-grade' comparison, gives information on how grade retainees developed in the year of retention. However, it usually draws a rather positive picture of the effects of grade retention, because students in the retained group received the subject matter twice. A solution to overcome this possible bias, is to use a 'same-age' approach where retained students are compared with their same-aged peers who are one year higher. If we would follow (only) this same-age approach, we would overestimate the impact of grade retention in the negative sense, because this kind of same-age comparisons are usually in favour of promoted students who were able to acquire more knowledge. The suggestion of several researchers is to use both kind of comparisons, and to keep in mind the possible bias related to every method (e.g. Bonvin et al., 2008; Goos et al., 2013). This was not possible with the LOSO data, because grade retainees were only measured one year later in Grade 8, 10 and 12, and there are no data available for promoted students in Grade 9 or Grade 11. If we wanted to follow a same-age approach, we would have needed measurements in Grade 9 and Grade 11 for promoted students, or for retained students (to have them both for promoted and retained students is not required). This imposes some restrictions on our conclusion, because we probably overestimate the benefits of grade retention (Allen et al., 2009). To draw valid conclusions on the effect of grade retention in secondary education, future research should take into account both approaches. Although it may seem very time-consuming to collect data at all these time-points, it is the only way to properly evaluate the effect of grade retention.

A fourth limitation is also related to Chapter 4, and concerns the so-called 'post-treatment covariates'. In this chapter, we matched students based on their propensity to be retained in Grade 8. The construction of the PS fully relies on the pre-treatment variables (Caliendo & Kopeinig, 2008; Rosenbaum & Rubin, 1983; Rubin, 2008), because variables included in the PS should be unaffected by the treatment. Based on this PS, we evaluated the effects of grade retention in the long-term. However, we should be aware of the fact that 'post-treatment covariates' (i.e. variables measured after grade retention) can also affect our outcome. This can be the case if, for example, students change track and some courses take a less

prominent place in the curriculum, and we believe this can (partially) explain the drop in achievement in Grade 12. One can think of several other variables which may influence long-term outcomes (for instance: retained in grade for the second time, although these students were not in our dataset, because they were lost to follow-up). These post-treatment variables should be included in further, more thorough analyses. Including the effect of post-treatment covariates is, in our opinion, an important future line of research.

A fifth limitation of our dissertation, is the operationalization of 'success' in Chapter 5. We defined success/failure in a very strict way, which may not be in accordance to the current reality in higher education. In our study, successful students were students who took all exams at the end of the first year, and was granted access to the next year. Unsuccessful students were students who (1) took all exams but failed several (even after a resit) or (2) did not take all exams at the end of the first year or (3) stopped before taking exams. Nowadays, higher education is more flexible than that, and students can choose which exam they take up in every year (with some restrictions due to prerequisites for some courses). As a result, students can also be successful if they only took up 50% of the exams in the first year, and passed these exams. So ideally, we evaluate success by comparing the number of passed exams to the number of exams the student took. However, the LOSO-database does not provide this information, because this kind of flexibility was not possible at the time the LOSO-survey was administered. If a student was registered as 'failed' in a certain year in higher education, the only information provided was if the student took all the exams or only for a part of the exams, and if he failed or passed the exams. Still, this does not solve our problem, because there is no information on the number of exams a student took in the first place (and the number of exams that were successful). Apart from the conceptualization of 'success', there are some issues with the time-frame of 'success'. We defined 'long-term success' as successful in every year, up until the third year in higher education (corresponding to a 'Bachelor's degree'). Following from this time-frame, students who obtained their Bachelor's degree with one (or more) year(s) delay are considered as unsuccessful. This rather short time-frame was the only – reliable – option with the available data, since we only have data on the success/failure until 3

years after graduating from secondary education. Afterwards, the information is not available for every student. Future research should have a longer time perspective, and should also take into account diploma's obtained after the 'normal' period. It is not unlikely that students lag one year behind in higher education, because the choice of a major in higher education is a very complex task, where a lot of different processes should be taken into account (Pinxten, De Fraine, Van Den Noortgate, Van Damme, Vanlaar & Boonen, 2013). Pinxten et al. also focussed on the decision after the first year in higher education, and concluded that more than half of the students failed in the first year (52%), but only 3% of these 'failures' left higher education. The other students opted for retaking the same major (51%) or changed over to another major (46%).

A sixth limitation is also related to the conceptualisation of a dependent variable: the variable measuring early school leaving. In Chapter 2 and Chapter 3, we concluded that early school leaving starts from Grade 9. At first sight, dropout from Grade 9 may seem odd and very early, but we see two possible explanations for this 'very early' dropout. First, one should keep in mind that we defined dropouts as students who leave fulltime secondary education. Consequently, students transferring to part-time education are also seen as 'early school leavers'. This can be seen as a limitation of our study, because in this part-time education, it is still possible to obtain an alternative qualification. However, we had a good reason to opt for our strict conceptualisation: comparability with previous research, and in that sense, we followed the reasoning of Bowers (2010). Bowers states that in the US, the options for students who do not wish to graduate (on time) are many and classification of dropout/graduate is sometimes very difficult. One possible alternative that students have in the US, is to work towards a General Educational Development credential (GED). In the public opinion, this GED graduates are considered as equivalent to students who received a 'traditional' high school diploma (and therefore, the abbreviation 'GED' is often translated as 'General Equivalency Degree') (Heckman, Humphries, & Mader, 2010). However, this is not true at all, because several studies noticed that a GED does not guarantee the same chances on the labour market as a 'traditional' high school diploma (Heckman et al., 2010; Tyler, 2003). Because of this negative outcome, Bowers decided to label

'GED' graduates also as dropouts. In our studies, we followed a similar reasoning for students who transferred to part-time education. De Rick (2006) concluded that – even for students leaving the alternative system with a qualification – those part-time alternatives do not guarantee the same chances on the labour market. It seems that part-time graduates are not the most attractive group of workers for an employee, because they lack the right skills. In our opinion, this can justify our conceptualisation of early school leaving. A second reason for 'early' dropout, is that grade retention is a strong predictor of early school leaving and a lot of dropouts are older than their classmates. Of course, this grade retention is not the main explanation for the dropout-rate in Grade 9 (if that would be the case, students would lag 5 years behind), but can be an explanation for the strong increase of the dropout-rate in the subsequent grades. As Van Landeghem and Van Damme (2011a) illustrated, students who experienced grade retention several times, were unlikely to successfully finish secondary education.

PRACTICAL IMPLICATIONS

Our results also have some practical implications for different levels of education. In this section, we reformulate our results in terms of (practical) advices for teachers and educational policy advisors. We first discuss the effect of changing track or repeating a grade after a B-certificate. Afterwards, we formulate some advice on (mainly) the prevention and intervention of early school leaving. The effects of grade retention will be interwoven in these discussions.

The options after a B-certificate: towards a binding advice?

In Chapter 4, we saw that students who opted for retention after receiving a B-certificate are worse off in the long-run regarding their achievement. When a student receives a B-certificate, he can choose between two options: repeating the same grade in the same track, or changing over to another (lower) track. But is it a good idea to keep the option open to remain in the same track? Or should the advice to change track be more a binding advice, ruling out the option of grade retention? We would opt for the latter, for two reasons. First, in our dataset, we matched students with comparable background characteristics, which resulted in a more

restricted sample. But, if we look at the results of previous studies where they used a more general sample and focussed specifically on the effect of socioeconomic background on educational choice, it was found that educational choice is strongly influenced by the socioeconomic status of the student's parents (Kloosterman & De Graaf, 2010; Spruyt et al., 2009). Both authors concluded that when the more socially privileged students are faced with the question of repeating a grade or changing track, they opt more for repeating the grade in the same track, compared to underprivileged students. This mechanism leads to a reproduction of social inequality: next to the effect of socioeconomic status on achievement, these underprivileged students end up more in the lower tracks because of their background. Indeed, as Boudon (1974) argues: socially underprivileged children do not only perform lower, they also make less ambitious choices concerning their educational career (mostly because their 'reference figure' (e.g. one of the parents) is less qualified, resulting in lower educational aspirations of the student). The second reason for making the advice more binding, is in the rationale behind the different certificates. What message does a teacher wants to give to the student, when the student receives a B-certificate? Does it mean that a student is able to handle the content of the specific track but did not master some aspects of it, and granting an extra year gives him the opportunity to master these aspects? Or does it mean that the track-specific content in general is probably too difficult for the student? If the latter is the main principle for granting a B-certificate, it would be better to make the advice to change track, binding. Moreover, if it is not necessary to make a choice, if there is no option, then this can also (partially) rule out the negative effects of socioeconomic status on educational choices. A forced change of track comes with some problems regarding the perception of certain tracks. When a student receives the (binding) advice to change over to another track, the implicit message can be that the current track is too difficult and a less difficult track is a better option, which, in turn creates a hierarchy between tracks. This encourages an 'aim-high-fall-deep' reasoning with parents, who like to see their children in the academic track. When this academic track is too demanding for the student, then he can change over to technical or vocational education. That way, (especially) choosing for the vocational track will never be a positive choice, but always a

forced choice because of low performance. For this reason, Pattyn et al. (2012) propose to dispense with the B-certificates.

Early school leaving: how do we keep students in school?

In order to tackle the problem of early school leaving, the European Commission emphasizes the importance of a comprehensive policy and proposes a Prevention-Intervention-Compensation Framework (PICF, European Commission, 2011). Actions in the prevention part of the framework aim at creating a stimulating learning climate for all students, and seek to tackle early school leaving even before the first warning signals of dropout come up. The intervention comes into action when the first signals of early school leaving turn up, and compensation is about offering alternative routes to a qualification for those students who already left school unqualified. It would lead us too far to discuss all the possible actions in the PICF, so we only discuss those actions that are relevant in the light of this dissertation. For a more complete overview of the PICF in Flanders, we refer to other publications (e.g. Lamote & Van Damme, 2011; Lamote, Van Droogenbroeck, & Van Damme, 2012; Lamote, Van Landeghem et al., 2013).

As one of the main preventive actions, the European Commission recommends a systematic language support, especially in the early years of education. This is in line with Van Damme (2010) who states that 'until the age of 10, you learn to read. After the age of 10, you read to learn'. If a child is not able to read fluently by the end of primary school, it will almost certainly fall behind in secondary education, increasing the risk of early school leaving. Although it cannot be denied that language support is important, we believe that – by the end of primary school – it is not the only subject that matters. Based on the conclusions in Chapter 2 and Chapter 3, we found that the initial achievement in general is a strong predictor of early school leaving. The variable measuring the initial achievement was a combination of a score on a language test, a mathematics test and an intelligence test, all administered at the beginning of secondary school. Of course, we admit that a good language proficiency is the basis for learning other subjects, but at the very beginning of secondary education, a teacher should be aware of the strengths and weaknesses of every student. A second preventive practice is situated

at the system level of education. The European Commission recommends to substitute grade retention with a more targeted support, because of the implications of grade retention (mainly: it has not proven to be effective to close lacks of competence). This is in line with the conclusions in Chapter 4 and Chapter 5. Indeed, also in Flanders, grade retention seems to have more negative effects than parents (and educators) expect. So, what should be done? Because most of the time low achievement is the main reason to hold a student back, the main advice would be to monitor – from the very beginning – a student's progress, and to intervene when a student is falling behind. On the other hand, if a student is falling behind because of a wrong option choice, a good (binding) advice should be given to this student. Especially for the latter, we made clear in Chapter 4 that grade retention should be no option. If a student is clearly in the wrong track, there is no use in repeating the grade, and a good orientation seems to be the better option. This leads us to our third advice on the prevention of early school leaving: transparency and permeability of educational pathways. When a student has to change track, it should be clear where every (new) track leads to, and which professions and post-secondary options remain open. In Lamote et al. (2012) we also suggested the importance of these short- and long-term perspectives, but emphasized that a choice of track should not be reduced to a pure 'economical approach' (i.e. which diploma leads to the shortest time in unemployment, or even worse, which diploma leads to the job with the highest wage?), but should be well-considered decision, taking into account, for example, interest in a particular subject matter. In that sense, the choice for vocational education should be a positive choice, instead of the endpoint of forced choices. A reader may think that this is in contradiction with our plea for a binding advice after a B-certificate, implying 'forced choices'. However, we believe that, even with a binding advice, a student can still find a track in his line of interest but on a different level. For example: if a student receives a B-certificate in the economics-section, he can still switch over to the technical option of 'trade', which is comparable to the more academic option of economics. This reasoning reflects one of the main ideas of the proposed educational reform in Flanders. In this educational reform, students can choose (starting from Grade 9) between 5 'domains' of subjects: science & technology, language & culture, health & society, art & creation and economics & organisation. Within this domains, subjects are

taught on different levels: on a level which prepares for higher education, on a level preparing for entrance in the labour market and on a level which serves both outcomes (higher education and labour market). This so-called 'matrix' structure makes it possible to follow a subject of interest on different levels.

As for the intervention part of the PICF, we mainly focus on the construction of early warning systems. Based on the results of Chapter 3, it is clear that at the very beginning of secondary education we can already find a group of students with an increased risk of dropping out. Especially the group of low engaged students can already be identified at the beginning of Grade 7, but also the group of students starting at a higher level and following a decreasing pattern can be identified at a very early stage. Both groups share a common set of covariates predicting group membership; most of the time it are boys stemming from a low SES family with a history of grade retention or a start in the remedial class. If these indicators, together with engagement measures, are integrated in some form of early warning system, and scoring high/low on one of these indicators activates some kind of 'red blinking warning light', this can be the start of an intervention. Based on several meta-analyses (De Witte & Cabus, 2013; Dynarski, Clarke, Cobb, & Finn, 2008; Wilson, Tanner-Smith, Lipsey, Steinka-Fry, & Morrison, 2011), we would recommend the most effective intervention: assigning a mentor to students at-risk. This mentor is matched to a student and is responsible for addressing academic and social needs, and can also act as a bridge between school and the parents.

The compensation part of the PICF, together with several other possible prevention and intervention practices are beyond the scope of this dissertation. More information on these aspects can be found in other publications (De Witte et al., 2013; Lamote et al., 2011, Lamote et al., 2012; Lamote, Van Landeghem et al., 2013).

EPILOGUE

After 4 years of intensive study and analysis of early school leaving and grade retention, every answer we found came along with new questions. With this dissertation, we tried to answer some – at first sight very easy – questions: has grade retention positive or negative effects, and what predicts early school leaving and how can we tackle this? These questions turned out to be very complex questions, not only because of methodological reasons, but also because of the interpretation of the results. Based on our results, especially for grade retention, it is difficult to formulate a simple answer without any nuance. Each of our studies had limitations which also contributed to the need for nuance. The reader should consider this dissertation as only one puzzle piece of two complex phenomena.

Is there a take-home message from this dissertation? The first one would be to think of solutions in terms of long-term effects, instead of opting for quick fixes (see: grade retention). The second one would be a paradigm change from 'student as pure learner' to 'student as a full person' by taking into consideration several aspects of engagement and other non-cognitive aspects. The third, and final message would be to turn away from simple answers/analyses to complex questions, but always dig into the data with advanced methodologies. Results based on thorough analyses are the only reliable basis for educational improvement.

References

Alexander, K., Entwisle, D., & Dauber, S. (2003). *On the success of failure: A reassessment of the effects of retention in the primary grades*. Second Edition. New York: Cambridge University Press.

Alexander, K., Entwisle, D., & Horsey, C. (1997). From First Grade Forward: Early Foundations of High School Dropout. *Sociology of Education, 70,* 87-107.

Alexander, K., Entwisle, D., & Kabbani, N. (2001). The dropout process in life course perspective: Early risk factors at home and school. *Teachers College Record, 103*, 760–822.

Allen, C., Chen, Q., Willson, V., & Hughes, J. N. (2009). Quality of design moderates effects of grade retention on achievement: A meta-analytic, multi-level analysis. *Educational Evaluation and Policy Analysis, 31*, 480-499.

Allison, P. (1982). Discrete-time methods for the analysis of event histories. *Sociological Methodology, 13*, 61–98.

Archambault, I., Janosz, M., Morizot, J., & Pagani, L. (2009) Adolescent behavioral, affective, and cognitive engagement in school: Relationship to dropout. *Journal of School Health, 79*, 408–415.

Barber, J. S., Murphy, S. A., Axinn, W. G., & Maples, J. (2000). Discrete-time multilevel hazard analysis. *Sociological Methodology, 30,* 201-235.

Bates, D., Maechler, M., & Bolker, B. (2012). *lme4: Linear mixed-effects models using S4 classes*. R package version 0.999999-0.

Becker, H. S. (1963). Outsiders. Studies in the sociology of deviance. New York: Free Press.

Bills, D. B. (2003). Credentials, signals, and screens: explaining the relationship between schooling and job assignment. *Review of Educational Research, 73*, 441-449.

Bingham, G.E. & Okagaki, L. (2012). Ethnicity and Student Engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 315-342). New York, Springer.

Blondal, K. S., & Adalbjarnardottir, S. (2012). Student Disengagement in Relation to Excepted and Unexpected Educational Pathways. *Scandinavian Journal of Educational Research, 56*, 85-100.

Bonvin, P., Bless, G., & Schuepbach, M. (2008). Grade retention: decision-making and effects on learning as well as social and emotional development. *School Effectiveness and School Improvement, 19*, 1-19.

Boudon, R. (1974). Education, opportunity, and social inequality: Changing prospects in Western society. New York: Wiley.

Bowers, A. (2010). Grades and Graduation: A Longitudinal Risk Perspective to Identify Student Dropouts. *The Journal of Educational Research, 103*, 191–207.

Browne, W. J., Goldstein, H., Rasbash, J. (2001). Multiple membership multiple classification (MMMC) models. *Statistical Modelling, 1*, 103–124.

Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*, 31-72.

Chen, Q. (2012). The Impact of Ignoring a Level of Nesting Structure in Multilevel Mixture Model: A Monte Carlo Study. *Sage Open, 2*, 1-10.

Christenson, S.L., Reschly, A.L., & Wylie, C. (2012). *Handbook of Research on Student Engagement*. New York, Springer.

Creten, H., Van de Velde, V., Van Damme, J. & Verhaest, D. (2004). *De transitie van het initieel beroepsonderwijs naar de arbeidsmarkt met speciale aandacht voor de onderwijsverlaters [The transition from vocational education to the labour market with a special reference to unqualified school leavers]*, Leuven: HIVA.

Croninger, R.G., & Lee, V.E. (2001). Social capital and dropping out of high school: Benefits to at-risk students of teachers' support and guidance. *Teachers College Record, 103*, 548-581.

De Fraine, B. (2003). *Cognitieve en niet-cognitieve effecten van prestatiegericht klimaat, gemeenschapsgericht klimaat en groepssamenstelling in klassen en scholen. Verkenning met het multiniveaumodel [Cognitive and non-cognitive effects of academic climate, communitarian climate and group composition of classes and schools. Explorations with multilevel models]*. Dissertation, KU Leuven, Belgium.

De Rick, K. (2006).Werkervaring voor leerlingen uit de deeltijdse leersystemen: motieven en ervaringen van de werkgevers. Leuven: HIVA.

De Witte, K., & Cabus, S. (2013). Dropout prevention measures in the Netherlands, an evaluation. *Educational Review, 65*, 155-177.

De Witte, K., Nicaise, I., Lavrijsen, J., Van Landeghem, G., Lamote, C., & Van Damme, J. (2013). The impact of institutional context, education and labour market policies on early school leaving: a comparative analysis of EU countries. *European Journal of Education*, *48*, 331-345.

Dynarski, M., Clarke, L., Cobb, B., Finn, J., Rumberger, R., & Smink, J. (2008). *Dropout Prevention: A Practice Guide (NCEE 2008-4025).* Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sience, U.S. Department of Education.

EACEA Eurydice (2012) Key Data on Education in Europe 2012. Brussels: Eurydice.

Eccles, J. (1983). Expectancies, values and academic behaviors. In J. Spence (Ed.), *Achievement and achievement motives* (pp. 75-146). San Francisco, W. H. Freeman.

Eccles, J., & Wang, M. T. (2012). Part I commentary: so what is student engagement anyway? In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 133-145). New York, Springer.

Ehmke, T., Drechsel, B., & Carstensen, C. H. (2010). Effects of grade retention on achievement and self-concept in science and mathematics. *Studies in Educational Evaluation, 36*, 27-35.

European Commission (2010). *Europe 2020: a strategy for smart, sustainable and inclusive growth*. Brussels, European Commission.

European Commission (2011). *Reducing early school leaving. Staff working paper*. Brussels: European Commission.

EUROSTAT (2013). *Headline indicators EU 2020.* Retrieved from http://epp.eurostat.ec.europa.eu/portal/page/portal/europe_2020_indicators/headline_indicators.

Ferguson, P., Jimerson, S., & Dalton, M. (2001). Sorting out successful failures: Exploratory analyses of factors associated with academic and behavioral outcomes of retained students. *Psychology in the Schools, 38*, 327-341.

Festinger, L. A. (1954). A theory of social comparison processes. *Human Relations, 7*, 117-140.

Fielding, A., & Goldstein, H. (2006). *Cross-classified and multiple membership structures in multilevel models: an introduction and review.* University of Birmingham: Department for University Skills.

Fielding, A. (2002). Teaching Groups as Foci for Evaluating Performance in Cost-Effectiveness of GCE Advanced Level Provision: Some Practical Methodological Innovations. *School Effectiveness and School Improvement, 13*, 225-246.

Fine, J., & Davis, J. (2003). Grade retention and enrollment in post-secondary education. *Journal of School Psychology, 41*, 401-411.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 59,* 117–142.

Finn, J. D., & Zimmer, K. (2012). Student engagement and academic achievement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 97-132). New York, Springer.

Fredricks, J. A., Blumenfeld, P. C., Friedel, J., & Paris, A. H. (2005). School engagement. In K. A. Moore & L. Lippman (Eds.), *Conceptualizing and measuring indicators of positive development: What do children need to flourish* (pp. 305–321). New York: Kluwer Academic/Plenum Press.

Fredricks, J. A., Blumenfeld, P. C., Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research, 74,* 59-109.

Fredricks, J. A. & McColskey, W. (2012). The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-report Instruments. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 763-782). New York, Springer.

Fruehwirth, J. C., Navarro, S., & Takahashi, Y. (2011). *How the timing of grade retention affects outcomes: Identification and estimation of time-varying treatment effects*. Chicago: Human capital and economic opportunity working group.

Fullarton, S (2002). Student engagement with school : individual and school-level influences, LSAY Research Reports n.27 http://research.acer.edu.au/lsay_research/31.

Gesthuizen, M., De Graaf, P.M., & Kraaykamp, G. (2005). The changing family background of the low-educated in the Netherlands, socio-economic, cultural, and socio-demographic resources. *European Sociological Review, 21*, 441-452.

Goldschmidt, P., & Wang, J. (1999). When can schools affect dropout behavior? A longitudinal multilevel analysis. *American Educational Research Journal, 36*, 715-738.

Goldstein, H. (1997). Methods in school effectiveness research. *School Effectiveness and School Improvement, 8*, 369-395.

Goldstein, H. (2011). *Multilevel statistical models*. London: Arnold Publishers.

Goldstein, H., Burgess, S., & McConnell, B. (2007). Modelling the effect of pupil mobility on school differences in educational achievement. *Journal of the Royal Statistical Society, 170*, 941–954.

Goldstein, H., & Sammons, P. (1997). The Influence of Secondary and Junior Schools on Sixteen Year Examination Performance: A Cross-classified Multilevel Analysis. *School Effectiveness and School Improvement, 8*, 219-230.

Goos, M., Van Damme, J., Onghena, P., Petry, K.,& de Bilde, J. (2013). First-grade retention in the Flemish educational context: Effects on children's academic growth, psychosocial growth, and school career throughout primary education. *Journal of School Psychology, 51*, 323-347.

Grady, M. W., Beretvas, S. N. (2010). Incorporating student mobility in achievement growth modeling: A cross-classified multiple membership growth curve model. *Multivariate Behavioural Research, 45*, 393–419.

Hanushek, E. A. (2009). The economic value of education and cognitive skills. In G. Sykes, B. Schneider, & D. N. Plank (Eds.), *Handbook of Education Policy Research*, (pp. 39-56). New York: Routledge.

Hanushek, E. A., Woessmann, L. (2012). The economic benefit of educational reform in the European Union. *CESifo Economic Studies, 58*, 73–109.

Heckman, J. J., Humphries, J. E., & Mader, N. S. (2010) *The GED (Discussion paper series // Forschungsinstitut zur Zukunft der Arbeit, No.4975)*. http://hdl.handle.net/10419/36948.

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics, 24*, 411-482.

Hickman, G., Bartholomew, M., Mathwig, J., & Heinrich, R. (2008). Differential developmental pathways of high school dropouts and graduates. *Journal of Educational Research, 102*, 3–14.

Ho, D., Imai, K., King, G., & Stuart, E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*, 199-236.

Ho, D., Imai, K., King, G., & Stuart, E. (2011). Matchit: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*. http://gking.harvard.edu/matchit/.

Holmes, C. T. (1989). Grade level retention effects: A meta-analysis of research studies. In L. A. Shepard & M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp. 16-33). Philadelphia: Taylor & Francis.

Hong, G., & Raudenbush, S. (2005). Effects of kindergarten retention policy on children's cognitive growth in reading and mathematics. *Educational Evaluation and Policy Analysis, 27*, 205-224.

Hong, G. L., & Yu, B. (2008). Effects of kindergarten retention on children's social-emotional development: An application of propensity score method to multivariate, multilevel data. *Developmental Psychology, 44*, 407-421.

Horn, L., & Carroll, C. (1997). *Confronting the odds: students at risk and the pipeline to higher education*. Statistical analysis report NCES, Washington DC.

Hughes, J. N., & Kwok, O. (2007). Influence of student-teacher and parent-teacher relationships on lower achieving readers' engagement and achievement in the primary grades. *Journal of educational Psychology, 99*, 39-51.

Imbens, G. W. (2000). The role of propensity score in estimating dose-response functions. *Biometrika, 87*, 706-710.

Imbens, G. W., & Rubin, D. B. (to appear). *Causal inference in Statistics and Social Sciences*. New York: Cambridge University Press.

Jackson, G. B. (1975). The research evidence on the effects of grade retention. *Review of Educational Research, 45*, 613-635.

Janosz, M., Archambault, I., Morizot, J., & Pagani, L.S. (2008). School engagement trajectories and their differential predictive relations to dropout. *Journal of Social Issues, 64*, 21-40.

Jimerson, S. R. (1999). On the failure of failure: Examining the association between early grade retention and education and employment outcomes during late adolescence. *Journal of School Psychology, 37*, 243-272.

Jimerson, S. R. (2001). Meta-analysis of grade retention research: Implications for practice in the 21st century. *School Psychology Review, 30*, 420-437.

Jimerson, S. R., Anderson, G. E., & Whipple, A. D. (2002). Winning the battle and losing the war: Examining the relation between grade retention and dropping out of high school. *Psychology in the Schools, 39*, 441-457.

Jimerson, S., Carlson, E., Rotert, M., Egeland, B., & Sroufe, L. A. (1997). A prospective, longitudinal study of the correlates and consequences of early grade retention. *Journal of School Psychology, 35*, 3-25.

Jimerson, S., & Ferguson, P. (2007). A longitudinal study of grade retention: academic and behavioural outcomes of retained students through adolescence. *School Psychology Quarterly, 22*, 314-339.

Karweit, N. L. (1999). *Grade retention: Prevalence, timing, and effects* (CRESPAR Report 33). Baltimore (MD), The Johns Hopkins University, Center for Social Organization of Schools.

Kleinbaum, D. G., & Klein, M. (2005). *Survival analysis. A Self-Learning Text*. New York, Springer

Klem, A. M., & Connell, J. P. (2004). Relationships matter: Linking teacher support to student engagement and achievement. *Journal of School Health, 74*, 262–273.

Kloosterman, R., & de Graaf, P. M. (2010). Non-promotion or enrolment in a lower track? The influence of social background on choices in secondary education for three cohorts of Dutch pupils. *Oxford Review of Education, 36*, 363-384.

Kortering, L. J., & Christenson, S. (2009) Engaging Students in School and Learning: The Real Deal for School Completion. *Exceptionality, 17*, 5-15.

Lacante, M., Almaci, M., Van Esbroeck, R., Lens, W., & De Metsenaere, M. (2007). *Allochtonen in het hoger onderwijs. Factoren van studiekeuze en studiesucces bij allochtone eerstejaarsstudenten. Onderzoek in opdracht van het Ministerie van de Vlaamse Gemeenschap. [Immigrants in higher education. Factors of option choice and success for immigrant freshmen's. Research assigned by Ministry of the Flemish Community]*. Leuven/Brussel, KU Leuven & Vrije Universiteit Brussel.

Lamote, C., Pinxten, M., Van Den Noortgate, W., & Van Damme, J. (2013). *Is the cure worse than the disease? A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept.* Paper submitted for publication.

Lamote, C., & Van Damme, J. (2011). Iedereen gekwalificeerd? Een samenvatting van de kenmerken, oorzaken, gevolgen en aanpak van ongekwalificeerd uitstromen vanuit een Vlaams en Europees kader. Brussel: Vlaams Ministerie van Onderwijs en Vorming.

Lamote, C., Van Damme, J., Van Den Noortgate, W., Speybroeck, S., Boonen, T., & de Bilde, J. (2013). Dropout in secondary education: An application of a multilevel discrete-time hazard model accounting for school changes. *Quality and Quantity, 47*, 2425-2446.

Lamote, C., Van Droogenbroeck, I., & Van Damme, J. (2012). Eén leerling per uur valt uit en de klok tikt verder. Over vroegtijdig schoolverlaten en suggesties voor een brede aanpak. In: De Vry M. (Eds.), *Lerende Gemeente. Gids voor flankerende onderwijsbeleid* (pp 1-30). Brussel, Politeia.

Lamote, C., Van Landeghem, G., Blommaert, M., Nicaise, I., De Fraine, B., & Van Damme, J. (2013). Voortijdig schoolverlaten in Vlaanderen: een stand van zaken en een voorstel tot aanpak. In: Callens, M., Vanderleyden, L., & Noppe, J. (Eds.), *De sociale staat van Vlaanderen*. Brussel: Studiedienst van de Vlaamse Regering, *in press*.

Leckie, G. (2009). The complexity of school and neighbourhood effects and movements of pupils on school differences in models of educational achievement. *Journal of the Royal Statistical Society A, 172*, 537–554.

Lee, V., & Burkam, D. (2003). Dropping out of high school: The role of school organization and structure. *American Educational Research Journal, 40*, 353–393.

Li, L., & Hser, Y.I. (2011). On inclusion of covariates for class enumeration of growth mixture models. *Multivariate Behavioral Research, 46*, 266-302.

Lunceford, J., & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: a comparative study. *Statistics in Medicine, 23*, 2937-2960.

Luo, W., & Kwok, O. (2009). The impacts of ignoring a crossed factor in analyzing cross-classified data. *Multivariate Behavioural Research, 44*, 182-212.

Luyten, H., Bosker, R., Dekkers, H., & Derks, A. (2003). Dropout in the lower tracks of dutch secondary education: predictor variables and variation among schools. *School Effectiveness and School Improvement, 14*, 373–411.

Marks, G. (2007). Do schools matter for early school leaving? Individual and school influences in Australia. *School Effectiveness and School Improvement, 18*, 429–450.

Martin, A. J. (2011). Holding back and holding behind: Grade retention and students' non-academic and academic outcomes. *British Educational Research Journal, 37*, 739–764.

Marsh, H. W. (2007*). Self-concept theory, measurement and research into practice: The role of self-concept in educational psychology*. Leicester: British Psychological Society.

McCullough, M. E., Friedman, H. S., Enders, C. K., & Martin, L. R. (2009). Does devoutness delay death? Psychological investment in religion and its association with longevity in the Terman Sample. *Journal of Personality and Social Psychology, 97*, 866-882.

McLachlan, G. J., & Chang, S. U. (2004). Mixture modelling for cluster analysis. *Statistical Methods in Medical Research, 13*, 347-361.

Moser, S., West, S., & Hughes, J. (2012). Trajectories of math and reading achievement in low achieving children in elementary school: effects of early and later retention in grade. *Journal of Educational Psychology, 104*, 603-621.

Muthén, B. O. (2004) Latent variable analysis: Growth mixture modeling and related techniques for longitudinal data. In D. Kaplan (Ed.) *Handbook of quantitative methodology for the social sciences* (pp 345-368). Newbury Park, Sage.

Muthén, B. O., Brown, C. H., Masyn, K., Jo, B., Khoo, S. T., Yang, C. C., Wang, C. P., Kellam, S., Carlin, J., & Liao, J. (2002). General growth mixture modeling for randomized preventive interventions, *Biostatistics, 3*, 459-475.

Muthén, B. O., & Masyn, K. (2005). Discrete-Time Survival Mixture Analysis. *Journal of Educational and Behavioral Statistics, 30,* 27-58.

Muthén, L. K., & Muthén, B. O. (1998-2010). *Mplus User's Guide, Sixth Edition*. Los Angeles, Muthén & Muthén.

Nagin, D. S. (1999). Analyzing Developmental Trajectories: A Semi-parametric, Group-based Approach. *Psychological Methods, 4*, 139-177.

Nylund, K. L., Asparouhov, T. & Muthen, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling. A Monte Carlo simulation study. *Structural Equation Modelling, 14*, 535-569.

OECD (2013). *Education at a Glance 2013: OECD Indicators*. Paris, OECD Publishing. http://dx.doi.org/10.1787/eag-2013-en.

Opdenakker, M. C., Maulana, R. & den Brok, P. (2012) Teacher-student interpersonal relationships and academic motivation within one school year: developmental changes and linkage. *School Effectiveness and School Improvement, 23*, 95-119.

Opdenakker, M. C., & Van Damme, J. (2005). Differences between secondary schools: study about school context, group composition, school practice, and school effects with special attention to public and Catholic schools and types of schools. *School Effectiveness and School Improvement, 17*, 87-117.

Ou, S., & Reynolds, A. (2010). Grade retention, postsecondary education, and public aid receipt. *Educational Evaluation and Policy Analysis, 32*, 118-139.

Pampel, F. C. (2000). *Logistic Regression: A Primer. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-132*. Thousand Oaks (CA), Sage.

Parker, P. D., Marsh, H. W., Lüdtke, O., & Trautwein, U. (2013). Differential school contextual effects for math and English: integrating the big-fish-little-pond effect and the internal/external frame of reference. *Learning and Instruction, 23*, 78-89.

Pattanayak, C., Rubin, D., & Zell, E. (2011). Propensity score methods for creating covariate balance in observational studies. *Revista Espanola de Cardiologia, 64*, 897-903.

Pattyn, B., Colpin, H., De Cock, M., De Fraine, B., Dehaene, W., Groenez, S., Henkens, B., Nicaise, I., Masschelein, J., Ooghe, E., Raymaekers, B., Simons, M., Van Damme, J., & Verschueren, K. (2012*). Hervormingen in het secundair onderwijs. Metaforum visietekst 7*. Leuven: werkgroep Metaforum.

Pinxten, M., De Fraine, B., Van Damme, J., & Van Den Noortgate, W. (2013). *The reciprocal effects model and educational choice. A longitudinal study on academic self-concept, achievement and (post)secondary educational choice.* Unpublished dissertation, Leuven.

Pinxten, M., De Fraine, B., Van Den Noortgate, W., Van Damme, J., Vanlaar, G., & Boonen, T. (2013). "I choose so I am". A logistic analysis of major selection in university and successful completion of the first year. *Studies in Higher Education.*

Pustjens, H., Van de gaer, E., Van Damme, J., & Onghena, P. (2004). Effect of secondary schools on academic choices and on success in higher education. *School Effectiveness and School Improvement, 15*, 281-311.

Pustjens, H., Van de gaer, E., Van Damme, J., & Onghena, P. (2008). Curriculum choice and success in the first two grades of secondary education: students, classes, or schools? *School Effectiveness and School Improvement, 19*, 155-182.

R Core Team (2012). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rasbash, J., Steele, F., Browne, W., & Goldstein, H. (2009). *A user's guide to mlwin, v2.10 [Computer software manual]*. University of Bristol.

Raudenbush, S. W. (1993). A crossed random effects model for unbalanced data with applications in cross-sectional and longitudinal research. *Journal of Educational and Behavioral Statistics, 18*, 321-349.

Raudenbush, S. W. & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods*. Thousand Oaks: Sage Publications.

Ream, R. K., Rumberger, R. W. (2008). Student engagement, peer social capital, and school dropout among Mexican American and Non-Latino White students. *Sociology of Education, 81*, 109-139.

Reschly, A. L. & Christenson, S. L. (2012). Jingle, Jangle, and Conceptual Haziness: Evolution and Future Directions of the Engagement Construct. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp 3-20). New York, Springer.

Roderick, M. (1994). Grade retention and school dropout: Investigating the association. *American Educational Research Journal, 31*, 729–759.

Rodney, L. W., Crafter, B., Rodney, H. E., & Mupier, R. M. (1999). Variables contributing to grade retention among African American adolescent males. *Journal of Educational Research, 92,* 185-190.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.

Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association, 79*, 516-524.

Rubin, D. (2008). For objective causal inference: design trumps analysis. *The Annals of Applied Statistics, 2,* 808-840.

Rumberger, R. W. (1987). High-school dropouts, a review of issues and evidence. *Review of Educational Research, 57*, 101-121.

Rumberger, R. W. (1995). Dropping out of middle school: A multilevel analysis of students and schools. *American Educational Research Journal, 32*, 583–625.

Rumberger, R. W. (2001). Why students drop out of school and what can be done. In Dropouts in America: How severe is the problem? What do we know about intervention and prevention? Cambridge, Harvard University.

Rumberger, R. W. (2011). *Dropping out. Why students drop out of high school and what can be done about it.* Cambridge, Harvard University Press.

Rumberger, R. W., & Larson, K. (1998). Student mobility and the increased risk of high school dropout. *American Journal of Education, 107*, 1–35.

Rumberger, R. W., & Rotermund, S. (2012). The relationship between engagement and high school dropout. In S. L. Christenson, A. L. Reschly & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp 491-514). New York, Springer.

Rumberger, R. W., & Thomas, S. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education, 73*, 39–67.

Shepard, L. A. (1989). A review of research on kindergarten retention. In L. A. Shepard and M. L. Smith (Eds.), *Flunking grades: Research and policies on retention* (pp 64-78). Philadelphia: Palmer Press.

Singer, J., & Willett, J. (1991). Modeling the days of our lives: Using survival analysis when designing and analyzing longitudinal studies of duration and the timing of events. *Psychological Bulletin, 110*, 268-290.

Singer, J., & Willett, J. (2003). *Applied longitudinal data analysis: Modelling change and event occurrence.* Oxford, Oxford University Press.

Skinner, E. A., Wellborn, J. G., & Connell, J. P. (1990). What it takes to do well in school and whether I've got it: a process model of perceived control and children's engagement and achievement in school. *Journal of Educational Psychology, 82*, 22-32.

Smits, J., & Vorst, H. (1982) Schoolvragenlijst voortgezet onderwijs [School questionnaire secondary education]. Nijmegen, Berkhout.

Snijders, T., & Bosker, R. (1999). Multilevel Analysis. An introduction to basic and advanced multilevel modeling. London, Sage Publications.

Solga, H. (2002). 'Stigmatization by Negative Selection': Explaining Less-Educated People's Decreasing Employment Opportunities. *European Sociological Review, 18*, 159-178.

South, S., Haynie, D., & Bose, S. (2007). Student mobility and school dropout. *Social Science Research, 36*, 68–94.

Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B Statistical Methodology, 64*, 583–639.

Spruyt, B., Laurijssen, I., & Van Dorsselaer, Y. (2009). Kiezen en verliezen. Een analyse van de keuze na het krijgen van een B-attest in het Vlaams secundair onderwijs als een replicatie van Kloosterman en De Graaf. *Mens & Maatschappij, 84*, 279-299.

Stearns, E., Moller, S., Blau, J., & Potochnick, S. (2007). Staying back and dropping out: The relationship between grade retention and school dropout. *Sociology of Education, 80*, 210-240.

Stuart, E.A. (2010). Matching methods for causal inference: a review and a look forward. *Statistical Science, 25*, 1-21.

Stuart, E. A., & Green, K. M. (2008). Using full matching to estimate causal effects in non-experimental studies: Examining the relationship between adolescent marijuana use and adult outcomes. *Developmental Psychology, 44*, 395-406.

Teitler, J. O., & Weiss, C. C. (2000). Effects of neighborhood and school environments of transitions to first sexual intercourse. *Sociology of Education, 73*, 112-132.

Thoemmes, F. J., & Kim, E. S. (2011). A systematic review of propensity score methods in the social sciences. *Multivariate Behavioural Research, 46*, 90-118.

Tyler, J. H. (2003). Economic benefits of the GED: lessons from recent research. *Review of Educational Research, 73*, 369-403.

Uysal, S. D. (2010). *The effect of grade retention on school outcomes: An application of doubly robust estimation method*. Konstanz: University of Konstanz.

Van Damme, J. (2010). *Meer masters in de lagere school*. Klasse voor leerkrachten.

Van Damme, J., De Fraine, B., Van Landeghem, G., Opdenakker, M. C., & Onghena, P. (2002). A new study on educational effectiveness in secondary schools in Flanders: An introduction. *School Effectiveness and School Improvement, 13*, 383-397.

Van Damme, J., De Troy, A., Meyer, J., Minnaert, A., Lorent, G., Opdenakker, M.-C., & Verduyckt, P. (1997). *Succesvol doorstromen in de aanvangsjaren van het secundair onderwijs (Successful passing through the first years in secondary education)*. Leuven, Acco.

Van Damme, J., Meyer, J., De Troy, A., & Mertens, W. (2001). *Succesvol middelbaar onderwijs? Een antwoord van het LOSO-project*. Leuven: Acco.

Van de gaer, E., Pustjens, H., Van Damme, J., & De Munter, A. (2009). School engagement and language achievement: A longitudinal study of gender differences across secondary school. *Merrill-Palmer Quarterly, 55*, 374-405.

Van Landeghem, G., & Van Damme, J. (2011a). *Vroege schoolverlaters uit het voltijds beroepsonderwijs. Vertrekpunt voor een gedifferentieerde aanpak rapport nr. SSL/OD1/2011.39.* Leuven: Steunpunt Studie- en Schoolloopbanen (SSL).

Van Landeghem, G., & Van Damme, J. (2011b) *Vroege schoolverlaters in Vlaanderen. Evolutie van de ongekwalificeerde uitstroom tot 2009 - Kort verslag rapport nr. SSL/OD1/2011.46*. Leuven: Steunpunt Studie- en Schoolloopbanen (SSL).

Van Landeghem, G., De Fraine, B., Gielen, S., & Van Damme, J. (2013). *Vroege schoolverlaters in Vlaanderen in 2010. Indeling volgens locatie, opleidingsniveau van de moeder en moedertaal rapport nr. SSL/2013.05/1.2.0.* Leuven, Steunpunt Studie- en Schoolloopbanen (SSL).

VDAB (2013). Werkzoekende schoolverlaters in Vlaanderen [job seeking school leavers in Flanders]. Brussel, VDAB.

Verhofstadt, E., De Witte, H., & Omey, E. (2007). Higher educated workers: better jobs but less satisfied? *International Journal of Manpower, 28*, 135–151.

Wang, M. T., Willett, J. B. & Eccles, J. S. (2011). The assessment of school engagement: Examining dimensionality and measurement invariance by gender and race/ethnicity. *Journal of School Psychology, 49*, 465-480.

Wilson, S. J., Tanner-Smith, E. E., Lipsey, M. W., Steinka-Fry, K., & Morrison, J. (2011). *Dropout prevention and intervention programs: Effects on school completion and dropout among school aged children and youth.* Campbell Systematic Reviews 2011:8. DOI: 10.4073/csr.2011.8

Wouters, S., De Fraine, B., Colpin, H., Van Damme, J., & Verschueren, K. (2012). The effect of track changes on the development of academic self-concept in high school: A dynamic test of the Big-Fish-Little-Pond effect. *Journal of Educational Psychology, 104*, 793-805.

Wu, W., West, S. G., & Hughes, J. N. (2008). Effect of retention in first grade on children's achievement trajectories over 4 years: A piecewise growth analysis using propensity score matching. *Journal of Educational Psychology, 100*, 727-740.

Wu, W., West, S. G., Hughes, J. N. (2010). Effect of grade retention in first grade on psychosocial outcomes. *Journal of Educational Psychology, 102*, 135-152.

Xu, J. (2004). Family help and homework management in urban and rural secondary schools. *Teachers College Record, 106*, 1786-1803.

Yamamoto, K., & Byrnes, D. (1987). Primary children's ratings of the stressfulness of experiences. *Journal of Research in Childhood Education, 2*,117-121.

Appendices

APPENDIX 1: APPENDIX TO CHAPTER 3

Emotional engagement: Relationship with Teachers (10 Items; α = 0.88)

- I think that most of the teachers are very helpful when I have problems with school work.
- Some teachers are kinder to others than to me. (–)
- I feel at ease with most of the teachers.
- There are few teachers who help me well with my school work. (–)
- There are enough teachers who listen patiently when I ask something.
- I get on well with most of the teachers.
- There are few teachers who understand me. (–)
- Some teachers don't have the patience to explain things to me. (–)
- The teachers dislike me. (–)
- Most of the teachers treat me in a nice way.

Behavioural engagement: Attitude toward Homework (5 Items; α = 0.82)

- When I have homework, I put it off for as long as possible before I start (-)
- When I have homework, I start as soon as possible.
- I usually start doing my homework of my own accord.
- When I want to do something nice, I still complete my homework first.
- At home, I only start doing my homework when I am told to do so. (–)

Items that have a negative contribution to the scale are indicated by (–).

APPENDIX 2: APPENDIX TO CHAPTER 4

List of pre-treatment covariates

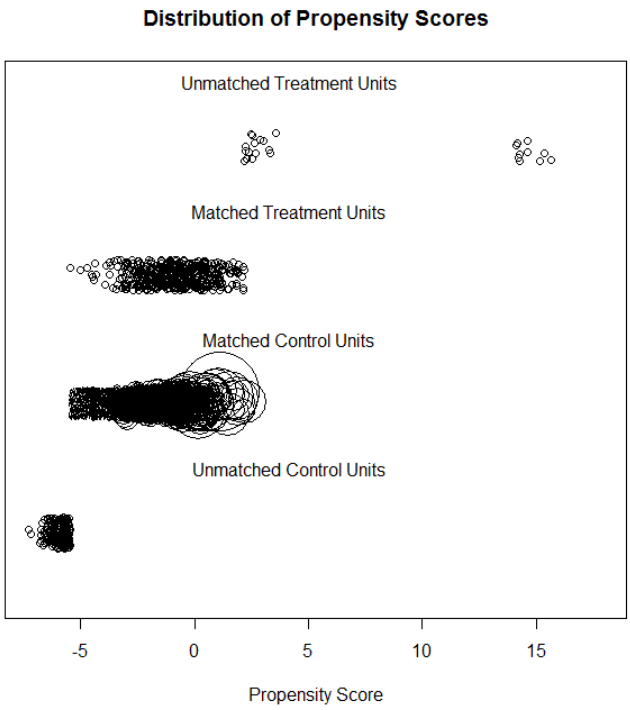
- Sex
- Math interest end of Grade 7 (teacher questionnaire)
- Socioeconomic status (SES)
- Which track in secondary education (teacher questionnaire primary education)
- Result end of primary school
- Parents have positive attitude towards school (teacher questionnaire primary education)
- Dutch achievement Grade 8, 1991-1992
- Intelligence test – reversed figures
- Academic self-concept Grade 8, 1991-1992
- In which track will I graduate?
- Dutch achievement begin Grade 7, 1990
- Student performed on a constant level (teacher questionnaire primary education)
- Match achievement begin Grade 7, 1990
- When I think about what I will do next year, I get a bad temper.
- Dutch achievement Grade 7, 1990-1991
- Interested in Dutch Grade 8 (teacher questionnaire)
- Match achievement Grade 7, 1990-1991
- Motivated Grade 8 (teacher questionnaire)
- Ethnicity
- Sex*Dutch achievement Grade 7
- Age at start secondary school
- Match achievement Grade 7*Age
- Attentiveness in the classroom 1990-1991
- SES*SES
- Attitude to homework 1990-1991
- Motivation towards learning tasks 1990-1991
- Attentiveness in the classroom 1991-1992
- Attitude to homework 1991-1992
- Expected result at the end of Grade 8 (teachers questionnaire)
- Expected certificate at the end of Grade 8 (teachers questionnaire)
- Involvement of parents (teacher questionnaire)
- Which group of Dutch achievement (teacher questionnaire)

APPENDIX 3: APPENDIX TO CHAPTER 5

Early retention – total sample – variables included in PS-matching

	Standardized mean difference	
Variable	Before matching	After matching
Results in primary school	-0.9985	-0.0067
Socio-economic status	-0.3729	-0.0239
Language achievement beginning of Grade 7	-0.8747	0.0637
Math achievement beginning of Grade 7	-0.8536	-0.0566
Language achievement end of Grade 7	-0.9858	0.1161
Math achievement end of Grade 7	-0.6489	0.066
Too old at start secondary school	0.1113	0.0376
Gender	-0.2967	0.069
Expected result (teacher questionnaire)	0.8381	0.0998
Follows class rules (teacher questionnaire)	0.7198	0.0348
Success rate	0.775	0.0427
Start in remedial class	0.1468	0.000
Effort in language class (teacher questionnaire)	-0.7987	0.0373
Positive attitude of parents towards school (primary school teacher questionnaire)	0.5231	0.0285
Interest in math (teacher questionnaire)	-0.8631	0.1016
Able to follow the lessons (teacher questionnaire)	0.9745	0.0654
Interest in numeric tasks	-0.0575	0.0187
Got help of parents in choosing a track in secondary education	0.0504	0.102
Teacher-student relations	-0.3473	0.0814
When I think about what I will do next year, it worries me.	-0.0309	-0.0521
When I think about what I will do next year, it tried to look at it from different angles.	-0.073	0.0473
Intelligence test – figures	-0.2833	0.0249
How long will I attend school?	-0.2873	-0.0008

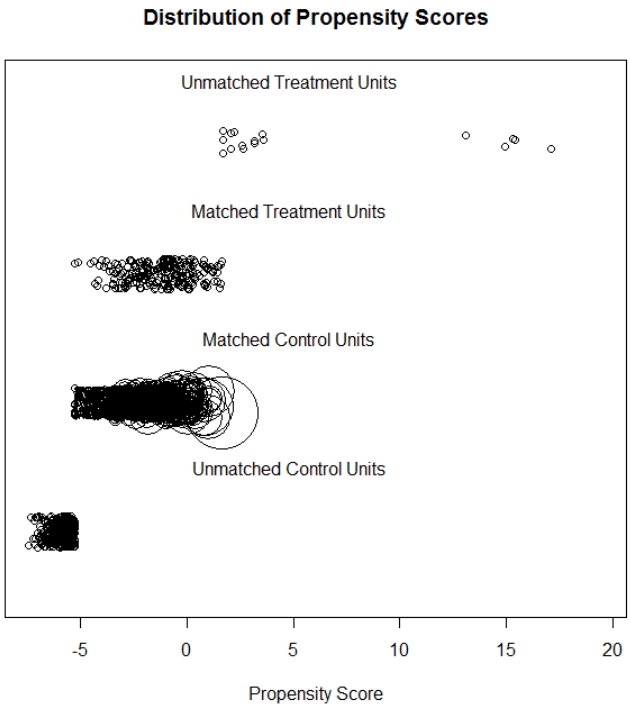
Early retention – total sample – Overlap



Early retention – higher education – variables included in PS-matching

	Standardized mean difference	
Variable	Before matching	After matching
Results in primary school	-1.0171	-0.0703
Socio-economic status	-0.2259	0.0622
Language achievement beginning of Grade 7	-0.8651	-0.064
Math achievement beginning of Grade 7	-0.8549	0.1508
Language achievement end of Grade 7	-0.9538	-0.0563
Math achievement end of Grade 7	-0.6786	-0.0647
Too old at start secondary school	0.105	0.0133
Gender	-0.2225	0.0096
Expected result (teacher questionnaire)	0.8071	0.0271
Success rate	0.859	-0.0316
Start in remedial class	0.14	0
Interest in language class (teacher questionnaire)	-0.7376	-0.0034
Teacher-student relations	-0.4015	0.0488
When I think about what I will do next year, I prefer to think about more nice things.	0.1419	0.0279
Believes in own capacities (teacher questionnaire)	0.7661	0.047
Effort in math class (teacher questionnaire)	-0.8269	0.0071
Language performance (primary school teacher questionnaire)	-0.8505	-0.0833
Positive attitude of parents towards school (primary school teacher questionnaire)	0.5234	-0.0567

Early retention – Higher education – Overlap

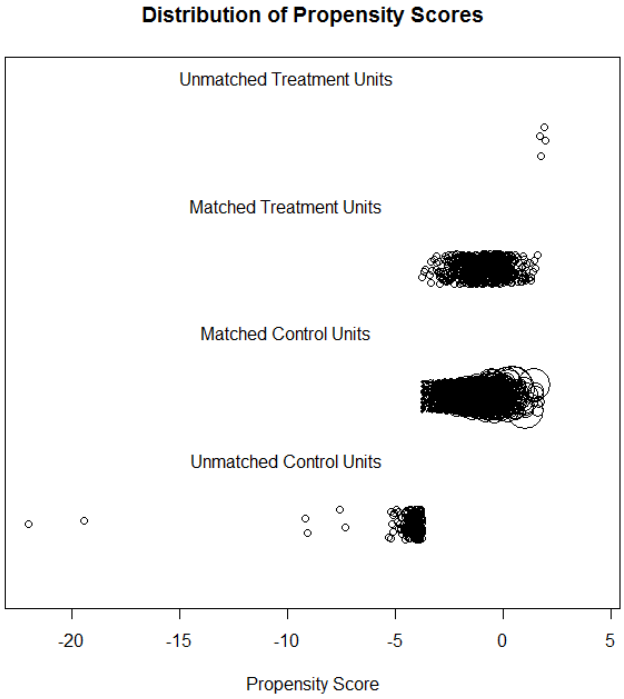


Late retention – total sample – variables included in PS-matching

	Standardized mean difference	
Variable	Before matching	After matching
Language achievement end of Grade 10	-0.6345	0.016
Math achievement end of Grade 10	-0.5034	-0.016
Results in primary school	-0.5251	0.0071
Socio-economic status	-0.312	0.0047
Language achievement beginning of Grade 7	-0.6165	-0.0447
Math achievement beginning of Grade 7	-0.5139	-0.0378
Language achievement end of Grade 7	-0.6499	-0.0169
Math achievement end of Grade 7	-0.411	-0.0084
Math achievement end of Grade 8	-0.3876	0.0723
Language achievement end of Grade 8	-0.634	-0.0199
Too old at start secondary school	0.1028	0.0229
Gender	-0.376	-0.0572
Motivation (teacher questionnaire Grade 8)	0.6435	0.0002
Which position does this student takes compared to the class (teacher questionnaire)	0.6096	0.0281
Homework attitude	-0.2918	0.0387
Special education in primary education	-0.3059	-0.0061
Effort in math class (teacher questionnaire Grade 8)	-0.6198	-0.0033
How long will I attend school?	-0.1312	0.0169
Language performance (primary school teacher questionnaire)	-0.4213	-0.008
Motivation (primary school teacher questionnaire)	0.4567	-0.0475
Positive attitude of parents towards school (primary school teacher questionnaire)	0.1534	-0.0269
When I think about what I will do next year, it tried to look at it from different angles.	0.0673	-0.1015
Importance of the choice of option	-0.0042	0.0142

Frequency of absence due to illness (primary school teacher questionnaire)	0.0173	-0.0608
This student could follow the lessons very well (primary school teacher questionnaire)	0.3389	0.0244
This student could work independently (primary school teacher questionnaire)	0.3511	0.0093
Doubts on the choice of secondary school	-0.0002	0.0052
Is this school new for you?	-0.039	0.0395
Intelligence test – synonyms	-0.2377	-0.0144
Importance of effort for study	0.1669	-0.0204
Do you think you are able to succeed in Grade 7?	0.0528	0.0041

Late retention – total sample – Overlap

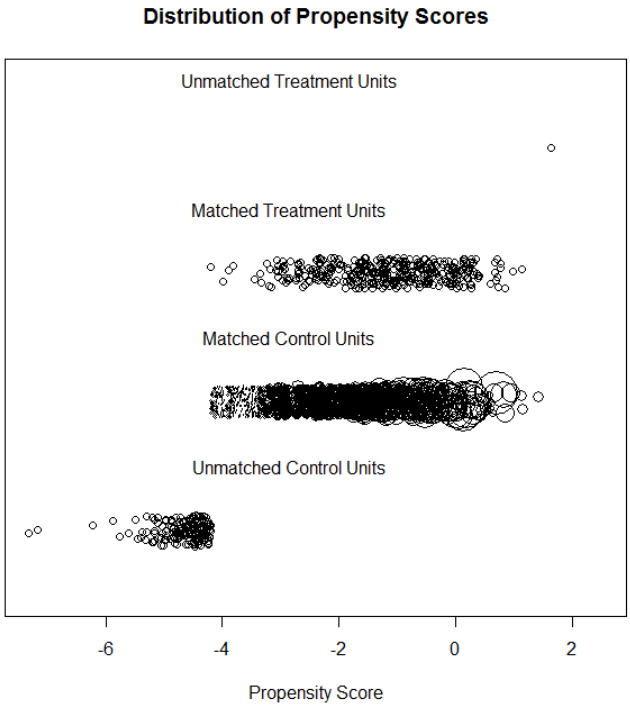


Late – higher education group – variables included in PS-matching

	Standardized mean difference	
Variable	Before matching	After matching
Language achievement end of Grade 10	-0.6558	-0.0118
Math achievement end of Grade 10	-0.4931	0.0046
Results in primary school	-0.4888	0.0811
Socio-economic status	-0.2147	0.0125
Language achievement beginning of Grade 7	-0.654	0.0791
Math achievement beginning of Grade 7	-0.539	0.1108
Language achievement end of Grade 7	-0.6797	0.0314
Math achievement end of Grade 7	-0.4498	0.1145
Math achievement end of Grade 8	-0.3984	0.0702
Language achievement end of Grade 8	-0.682	0.1315
Too old at start secondary school	0.0934	-0.0564
Gender	-0.3747	-0.026
Motivation (teacher questionnaire Grade 8)	0.5813	0.0469
Homework attitude	-0.2895	-0.0163
Which position does this student takes compared to the class (teacher questionnaire)	0.5982	0.0553
Importance of the choice of option	-0.068	-0.023
When I think about what I will do next year, it tried to look at it from different angles.	0.1069	-0.0264
How long will I attend school?	-0.0719	0.0602
Effort in math class (teacher questionnaire Grade 8)	-0.6156	-0.0288
Special education in primary education	-0.1684	-0.0081
Language performance (primary school teacher questionnaire)	-0.407	0.0594
Motivation (primary school teacher questionnaire)	0.4265	-0.0161
Positive attitude of parents towards school (primary school teacher questionnaire)	0.1081	0.0042
Doubts on the choice of secondary school	0.0361	-0.0532

Could your grades be better?	-0.3467	0.0159
Interest in art	-0.2685	-0.0027
In the search for a good track, I took all the pros and cons into account	0.1821	-0.0493
Do you think you are able to succeed in Grade 7?	0.0349	-0.0243

Late retention – higher education – Overlap



APPENDIX 4: FULL PUBLICATION LIST

Articles in internationally reviewed academic journals

Depaepe, F., Lamote, C., Vanlaar, G., Verhaeghe, J., Verschaffel, L., Van Damme, J. (2014). Leerlingpercepties en wiskundeprestaties: Een grootschalig onderzoek in het Vlaamse basisonderwijs naar het wiskundig zelfconcept, de inzet voor wiskunde en het schoolwelbevinden van leerlingen in relatie tot hun wiskundig functioneren. *Pedagogische Studiën, 91* (in press).

De Witte, K., Nicaise, I., Lavrijsen, J., Van Landeghem, G., Lamote, C., Van Damme, J. (2013). The impact of institutional context, education and labour market policies on early school leaving: a comparative analysis of EU countries. *European Journal of Education, 48*(3), 331-345.

Lamote, C., Van Damme, J., Van Den Noortgate, W., Speybroeck, S., Boonen, T., de Bilde, J. (2013). Dropout in secondary education: An application of a multilevel discrete-time hazard model accounting for school changes. *Quality and Quantity, 47* (5), 2425-2446.

Boonen, T., Speybroeck, S., de Bilde, J., Lamote, C., Van Damme, J., Onghena, P. (2013). Does it matter who your schoolmates are? An investigation of the association between school composition, school processes and mathematics achievement. *British Educational Research Journal, in press*.

Lamote, C., Speybroeck, S., Van Den Noortgate, W., Van Damme, J. (2013). Different pathways towards dropout: the role of engagement in early school leaving. *Oxford Review of Education, 39* (6).

de Bilde, J., Van Damme, J., Lamote, C., De Fraine, B. (2013). Can alternative education increase children's early school engagement? A longitudinal study from kindergarten to third grade. *School Effectiveness and School Improvement, 24* (2), 212-233.

Speybroeck, S., Kuppens, S., Van Damme, J., Van Petegem, P., Lamote, C., Boonen, T., de Bilde, J. (2012). The Role of Teachers' Expectations in the Association between Children's SES and Performance in Kindergarten: A Moderated Mediation Analysis. *PLoS One, 7* (1), 1-8.

Lamote, C., Engels, N. (2010). The development of student teachers' professional identity. *European journal of teacher education, 33* (1), 3-18.

Articles submitted for publication in internationally reviewed academic journals

Lamote, C., Pinxten, M., Van Den Noortgate, W., & Van Damme, J. (2013). *Is the cure worse than the disease? A longitudinal study on the effect of grade retention in secondary education on achievement and academic self-concept.*

Lamote, C., Van Den Noortgate, W., & Van Damme, J. (2013). *When you have beaten the odds: success of grade retainees in higher education.*

Articles in other academic journals

Burman, G., Lamote, C., Hannes, K., Van Damme, J. (2013). Waarom verlaten jongeren vroegtijdig het secundair onderwijs? *Impuls voor Onderwijsbegeleiding, 43* (3), 131-138.

Lamote, C., Van Den Noortgate, W., Van Damme, J. (2013). Vroegtijdig schoolverlaten in Vlaanderen: een stand van zaken en de rol van betrokkenheid. *Impuls voor Onderwijsbegeleiding, 43* (3), 122-130.

Articles in other professionally oriented journals

Goos, M., Belfi, B., Lamote, C., Van Damme, J. (2010). Zittenblijven op jonge leeftijd: 1 stap achteruit en vervolgens 2 stappen vooruit?. *Welwijs: Wisselwerking Onderwijs en Welzijnswerk, 21* (2), 26-29.

Article in other academic book

Lamote, C., Van Landeghem, G., Blommaert, M., Nicaise, I., De Fraine, B., & Van Damme, J. (2013). Voortijdig schoolverlaten in Vlaanderen: een stand van zaken en een voorstel tot aanpak. In: Callens, M., Vanderleyden, L., & Noppe, J. (Eds.), *De sociale staat van Vlaanderen.* Brussel: Studiedienst van de Vlaamse Regering, in press.

Lamote, C., Van Droogenbroeck, I., Van Damme, J. (2012). Eén leerling per uur valt uit en de klok tikt verder. Over vroegtijdig schoolverlaten en suggesties voor een brede aanpak. In: De Vry M. (Eds.), *Lerende Gemeente. Gids voor flankerende onderwijsbeleid.* Politeia, 1-30.

Lamote, C., Van Damme, J. (2011). Ongekwalificeerde uitstroom in Vlaanderen. In: Mahieu P. (Eds.), *School en samenleving*. Mechelen: Plantyn.

Meeting abstracts, presented at international scientific conferences

Depaepe, F., Lamote, C., Vanlaar, G., Verhaeghe, J., Verschaffel, L., Van Damme, J. (2013). Students' Affect and their Mathematical Learning Outcomes: A Longitudinal Study in Grade 4 and 5. AERA. San Francisco, 27 April - 1 May 2013.

Vanlaar, G., Denies, K., Pinxten, M., Vandecandelaere, M., Lamote, C. (2012). How to improve reading comprehension in high-risk students. AERA Annual meeting. Vancouver, BC, 13-17 April 2012.

Lamote, C., Van Den Noortgate, W., Speybroeck, S., Van Damme, J. (2012). Do Different Trajectories of Engagement lead to a Different Chance to Dropout?. European Conference on Educational Research (ECER). Cadiz, 18-21 September 2012.

Lamote, C., Speybroeck, S., Van Den Noortgate, W., Van Damme, J. (2012). Different pathways towards dropout; the role of engagement in early school leaving. EARLI SIG 18. Zürich, 29 - 31 August 2012.

Speybroeck, S., Kuppens, S., Lamote, C., Van Droogenbroeck, I. (2012). Child and Teacher Characteristics associated with Teachers' Expectations in Kindergarten: A Multilevel Analysis. AERA Annual meeting. Vancouver, BC, 13-17 April 2012.

Lamote, C., Van Damme, J., Van Den Noortgate, W. (2011). Dropout in Secondary Education: A Multilevel Discrete-time Hazard Model, Accounting for School Change. Biennial conference of the European Association for Research on Learning and Instruction (EARLI). Exeter, 30 August - 03 September 2011.

Meeting abstracts, presented at other scientific conferences and symposia

Lamote, C., & Van Damme, J. (2013). Verbondenheid met de school en vroegtijdig schoolverlaten: een verkenning van de relatie. Studiedag 'Voortijdig schoolverlaten: uit de tijd?', Studiedag van het Steunpunt Studie- en Schoolloopbanen. Leuven, 5 juni 2013.

Lamote, C., Van Damme, J. (2011). Vroegtijdig schoolverlaten in Vlaanderen. Van oorzaak tot aanpak. Seminarie Vroegtijdig Schoolverlaten Departement Onderwijs en Vorming. Brussel, 31 mei 2011.

Lamote, C., Van Damme, J., Van Den Noortgate, W. (2011). Vroegtijdig schoolverlaten in Vlaanderen, een kwestie van de school en de betrokkenheid?. Conferentie van het Steunpunt Studie- en Schoolloopbanen. Leuven, 24-25 februari 2011.

Lamote, C., Van Damme, J. (2011). Vroegtijdig schoolverlaten in Vlaanderen. Samen tot aan de meet: Lezingen en workshops. Antwerpen, 30 September 2011.

Lamote, C., Van Damme, J., Van Den Noortgate, W. (2011). Vroegtijdig schoolverlaten in secundair onderwijs; een multiniveau discrete-time survival analyse met kruisclassificatie. Onderwijs Research Dagen. Maastricht, 8-10 June 2011.

Internal report

Blommaert, M., Lamote, C., Meyer, J., & Van Damme, J. (2014). *De lange-termijneffecten van vroegtijdig schoolverlaten in Vlaanderen (werktitel)*. Leuven: Centrum voor Onderwijseffectiviteit en –evaluatie, KU Leuven.

Nicaise, I., De Witte, K., Lavrijsen, J., Lamote, C., Van Landeghem, G. (2013). *Towards a basic qualification for all in the EU – a social, educational and economic agenda*. Policy paper by KU Leuven Euroforum, 1-27 pp: Metaforum KU Leuven.

External reports

Nicaise, I., De Witte, K., Lamote, C., Lavrijsen, J., Van Landeghem, G. (2013). *Towards a Basic Qualification for All in the EU: A Social, Educational and Economic Agenda*. Euroforum policy paper 9, 1-27 pp.

Lamote, C., Van Damme, J. (2011). *Iedereen gekwalificeerd? Een samenvatting van de kenmerken, oorzaken, gevolgen en aanpak van ongekwalificeerd uitstromen vanuit een Vlaams en Europees kader*. Brussel: Vlaams Ministerie van Onderwijs en Vorming.

[i] Throughout this thesis, the terms 'early school leaving', 'dropout', 'unqualified school leaving' will be used interchangeably.

[ii] In Flanders, when students proceed through school without delay, they are 18 years old by the end of the 6th year of secondary education. This is the end of compulsory education

[iii] This is the definition of early school leaving of the Policy Research Centre on Educational and School Careers (*Steunpunt studie- en schoolloopbanen (SSL)).*

[iv] However, the term 'Bachelor's degree' can be misleading, because in Flanders, the Bachelor-Master structure has only been introduced since 2004-2005. Students in our dataset were in the third year at the earliest in 1999, so finalising the third year did not lead to a Bachelor's degree as we know it nowadays, especially not for students in the third year of an academic higher education college or a university since they did not receive a degree after the third year at all.

[v] For all the results reported in the text or in tables, the following legend applies: * $p < .05$, ** $p < .01$, *** $p < .001$, also when is not mentioned explicitly (often because of lay-out reasons).

[vi] In Flanders, part-time education is not an attractive alternative. Students transferring to this part-time education often do so because they did not reach the legal age to leave education (= 18 years). The majority of the students transferring to this part-time education alternative do not obtain the final certificate from this part-time education either. Therefore, we consider students transferring to this part-time alternative as 'real-dropouts'

[vii] Students are able to obtain a diploma of secondary education via alternative ways (and via part-time education), but the number of students who succeeds in obtaining this diploma is negligible.

[viii] In Table 9 and Table 12, we reported the log odds. The corresponding odds ratio can easily be computed by taking the antilog of the log odds.

[ix] Note that the intercept is a class-varying intercept with the 'High' category as a reference category (Muthén & Masyn, 2005). As we only have two categories, only one intercept is reported which can give the (wrong) impression that this

intercept is class-invariant. On the other hand, the covariates predicting survival are class-invariant. For more information on the format of reporting the outcomes of a DTSMA, see Muthén and Masyn (2005).

[x] For a full overview of the educational system in Flanders, we refer to p. 391-392 of Van Damme, De Fraine, Van Landeghem, Opdenakker and Onghena (2002).

[xi] The European indicator is not completely comparable to our LOSO-indicator, because a LOSO-student could enroll in higher education at the age of 18 in the school year 1996-1997. However, the 'oldest' European indicator for enrollment in higher education in Flanders dates back to 1999 (but is the number of students in tertiary education - as % of the population aged 20-24).

[xii] The list of variables is available upon request.

[xiii] Another possible use of the PS is to include this score in the regression as a covariate. By including it as a covariate, we are able to estimate a multilevel multinomial regression, but this technique is also criticized (Austin, Grootendorst, Normand, & Anderson, 2007). However, as a 'sensitivity check', we estimated such a model, which resulted in the same significant effects, compared to the weighted multinomial regression:
RET_E: $\beta_{RET\_E,HEC\_PROF} = -0.044, \text{SE} = 0.141, ns;\ \beta_{RET_E,HEC_{ACAD}} = 0.045, \text{SE} = 0.230, ns;\ \beta_{RET_E,SPEC} = -0.277, \text{SE} = 0.275, ns;\ \beta_{RET_E,WORK} = 0.624, \text{SE} = 0.145, p < .001$;
RET_L: $\beta_{RET\_L,HEC\_PROF} = 0.966, \text{SE} = 0.193, p < .001;\ \beta_{RET\_L,HEC\_ACAD} = 0.478, \text{SE} = 0.235, p = .042;\ \beta_{RET\_L,SPEC} = 0.612, \text{SE} = 0.235, p = .033;\ \beta_{RET\_L,WORK} = 1.544, \text{SE} = 0.222, p < .001$